AUGUST
1951
VOLUME 15
NUMBER 4

JOURNAL OF CONSULTING PSYCHOLOGY

AMERICAN PSYCHOLOGICAL ASSOCIATION

JOURNAL OF CONSULTING PSYCHOLOGY

EDITED BY LAURANCE F. SHAFFER

TEACHERS COLLEGE, COLUMBIA UNIVERSITY

ASSOCIATE EDITORS: EDGAR A. DOLL, *Devereux Schools, Devon, Pa.* WILLIAM A. HUNT, *Northwestern University.* E. LOWELL KELLY, *University of Michigan,* MORRIS KRUGMAN, *Board of Education, New York, N. Y.* BERTHA M. LUCKEY, *Board of Education, Cleveland, Ohio.* FRED MCKINNEY, *University of Missouri.* ANNE ROE, *New York, N. Y.* CARL R. ROGERS, *University of Chicago.* R. NEVITT SANFORD, *University of California.*

MARGARET K. HARLOW, *Managing Editor*

THE JOURNAL OF CONSULTING PSYCHOLOGY is the clinical journal of the American Psychological Association. It is devoted primarily to the publication of original research relevant to psychological diagnosis, psychotherapy and counseling, personality, and the dynamics of behavior. Case studies, theoretical contributions, descriptions of clinical techniques, and discussions of the professional problems of clinical psychology also appear in its pages.

EDITORIAL OFFICE

Manuscripts submitted for publication, books and tests for review, and correspondence on editorial matters should be sent to the Editor, Laurance F. Shaffer, 525 West 120th St., New York 27, N. Y.

Information for contributors: The important characteristics of a good article are the significance of its subject matter, the demonstrated soundness of its conclusions, its clarity, and its brevity. Clarity is improved by careful outlining, discriminating use of headings, and a simple, straightforward style of writing. For brevity, authors should strive to convey every thought without waste of words, and should avoid repetitions of methods, data or findings.

Some of the mechanical requirements for manuscripts are: Type manuscripts with wide margins, on white bond paper. Double space between *each and every line,* including references, footnotes, matter to be set in smaller type, and tables. Use only two types of headings: centered headings, and underlined paragraph headings. Type each table, however small, on a separate sheet, with Arabic number and title. Draw all figures in India ink in final form for reproduction, with draftsman's lettering, but type captions for figures on separate sheets. The *References* must follow the exact form used in APA journals. For further information on the details of preparing manuscripts, see:

> ANDERSON, J. E., AND VALENTINE, W. L. The preparation of articles for publication in the journals of the American Psychological Association. *Psychol. Bull.,* 1944, 14, 345-376.

There is a lag of about nine to twelve months between the acceptance of an article and its publication. Authors may secure immediate publication, however, by arranging for articles to be printed as extra pages, increasing the number of pages that subscribers receive. The cost to the author of such early publication is approximately $11.00 per printed page.

Authors receive fifty reprints gratis, and may order additional quantities when returning proof. Authors are charged one-half of the extra cost of tabular material and cuts, and the full cost of changes made on the proof.

BUSINESS OFFICE

Communications concerning subscriptions, change of address, claims for the nonreceipt of a number, advertising, and other business matters should be sent to the American Psychological Association, Inc., 1515 Massachusetts Ave., N. W., Washington 5, D. C.

The Journal of Consulting Psychology is published bi-monthly in February, April, June, August, October, and December, the yearly volume comprising approximately 480 pages. The subscription per year is $5.00, foreign $5.50, single number $1.00.

Copyright 1951 by the American Psychological Association, Inc.

Printed and mailed by the Dentan Printing Co., 23 West Colorado Ave., Colorado Springs, Colo. and entered as second-class matter at the Post Office at Colorado Springs, Colo. Acceptance for mailing at special rate of postage provided for in section 538, Act of February 25, 1925, authorized May 27, 1940.

JOURNAL OF CONSULTING PSYCHOLOGY

August, 1951 Vol. 15, No. 4

CONTENTS

Procedures and Methods in Supervising Clinical Interns: GLADYS L. ANDERSON - - 267

Supervision of Interns in a Community Guidance Center: JOSEPH E. BREWER - - 268

Supervision of Clinical Psychology Trainees in a Neuropsychiatric Hospital: A. L. CARP - 271

Training in Psychotherapy: O. H. MOWRER - - - - - - - - - - - - - 274

The Psychologist's Behavior as an Area of Research: SEYMOUR B. SARASON - - - 278

Oscillation and Progress in Clinical Psychology: MAUD A. MERRILL - - - - - 281

Symptom Patterns in a Group of Mental Hospital Patients: J. R. WITTENBORN - - 290

The Einstellung Test of Rigidity: Its Relation to Concreteness of Thinking: ABRAHAM S. LUCHINS - - - - - - - - - - - - - - - - - - - 303

The Performance of Young Schizophrenics and Young Normals on the Wisconsin Card Sorting Test: ELIZABETH T. FEY - - - - - - - - - - - - - 311

Nervous Symptoms among Military Personnel as Related to Age, Combat Experience, and Marital Status: RAYMOND G. KUHLEN - - - - - - - - - - - 320

The Wechsler-Bellevue and Descriptive Diagnosis: J. R. WITTENBORN AND JULES D. HOLZBERG - - - - - - - - - - - - - - - - - 325

The Wechsler Intelligence Scale for Children and the Revised Stanford-Binet: ARTHUR WEIDER, PAUL A. NOLLER, AND THEODORE A. SCHRAMM - - - - - 330

The Development of a Szondi Instability Score: HENRY P. DAVID AND WILLIAM RABINOWITZ - - - - - - - - - - - - - - - - - 334

The Influence of Set for Area on the Rorschach Test Results: LEONARD S. ABRAMSON - 337

Rorschach "Organic Signs" and Intelligence Level: WALLACE C. DIERS AND CLINTON C. BROWN - - - - - - - - - - - - - - - - - - 343

Performance on the Harrower-Erickson Multiple Choice Test of Patients with Spinal Cord Injuries: HAROLD H. MOSAK - - - - - - - - - - - - 346

MMPI Profiles in Patients with Low-Back Pain: LEO J. HANVIK - - - - - - - 350

New Books and Tests - 354

PROCEDURES AND METHODS IN SUPERVISING CLINICAL INTERNS[1]

GLADYS L. ANDERSON
MICHIGAN STATE COLLEGE

THE APA Committee on Graduate and Professional Training has recommended that one year of the four-year program of training in clinical psychology be devoted to an internship. The year may be one full year or it may be part-time work amounting to the equivalent of one year. It may all be spent in one clinic, or it may be divided, preferably to afford a variety of experiences in clinical work with male and female adults and with children and their parents.

In organizing this roundtable discussion on supervision of interns, an effort was made to present problems and viewpoints from a diversity of sources. Accordingly, the participants who were invited were able to draw on their experiences in the following areas: a community clinic, a neuropsychiatric hospital, university training in psychotherapy, and the examination of the reactions of the student himself as an area of research in supervision. Because supervision of interns in clinical training is currently a rather unstructured responsibility, it seem highly desirable that the papers presented at this roundtable be considered as a group, revealing in some degree the varied perspectives of the internship.

Supervision of students in the use of clinical tests is not a new responsibility for those engaged in the training of clinical psychologists. For years students have learned that the administration of a mental test must be carried out in a standard, prescribed, objective manner which can be learned in detail. Elaborate scoring manuals have been developed which must be followed rigorously in determining the nature and level of the subject's performance. Students have been carefully trained in the administration and scoring of psychological examinations. There has been an accepted core of basic materials which both the student and supervisor could consult at this level of training. Up to this point clinical training can be articulate, verbalized, and exact. Because it can be verbalized, it can be taught.

Beyond this point, however, there are few systematic patterns to follow. Here begins the clinical experience that is all too glibly referred to as an art. It is well known that some artists are better than others. Art, regarded in this sense, is merely unverbalized science. The "good" artist's clinical hunches may be valid, but because neither he nor anyone else is able to describe or narrate what he has done, his art remains an art. As soon as art rises to a level of communication, it can be taught, and it can next become a science. In the four papers which constitute the background for discussion at this roundtable on supervision are the verbalizations of many good hunches by clear-sighted clinical psychologists who believe their art can someday be taught.

[1]Chairman's introduction to a roundtable on the supervision of clinical interns. American Psychological Association, Pennsylvania State College, September, 1950.

Received November 30, 1950.

SUPERVISION OF INTERNS IN A COMMUNITY GUIDANCE CENTER[1]

JOSEPH E. BREWER
WICHITA GUIDANCE CENTER

THE intern training program that I will discuss has developed through the interaction of two forces. One has been the need for trained personnel—a need which was difficult to satisfy because psychologists in training had few opportunities to learn how to function in community psychological work. The Board of Directors of the Wichita Guidance Center realized early that it must devote some of its time and money to training personnel for replacement and expansion. The second force has been the thinking of the APA through its Committee on Graduate Training in Clinical Psychology. This thinking has evolved the general outlines of a professional training program that accords to clinical psychologists a highly responsible service function in collaboration with other professional personnel. The internship period has been structured as occurring in the third year of a four-year doctoral program.

The Wichita Guidance Center internship is designed to form an integral part of a doctoral training progam. It is further designed to give the broadest emphasis possible to this training by experience in a community psychological service, in sufficient extensity and intensity to prepare the trainee for professional competence at a high level of responsibility. We are not interested in developing technicians as such, but are interested in developing technical competence in clinicians. This requires that a full measure of responsibility be given the intern at as early a moment as possible in his year with us. It is important that judgments about the interns' personal integrity and sense of professional responsibility be made early and that assignment of cases and degree of responsibility permitted be gauged accordingly.

The operation of the Center has been described elsewhere. The average monthly caseload is 176. The age range of clients is from two-month infants to adults. The median age level is nine years. The staff to which the interns come consists of three senior psychologists, two junior psychologists, two social workers, two remedial teachers, and two part-time psychiatrists. A staff of this size can handle four interns.

The formal structure of the supervision received by the interns consists of (1) weekly individual conferences with the supervisor of the training program, (2) weekly group meetings with the supervisor, (3) weekly group meetings with each of the psychiatrists, (4) one to three staff case conferences per week, and (5) conferences by appointment with other staff members regarding individual cases. The interns are responsible to a supervisor for their general conduct, and for their program as a whole. Their clinical services are in the beginning under the direct control of a senior staff member who carries case responsibility. Later in their year they may receive only administrative supervision as would any other junior staff member.

The introduction of the intern to the services performed by the Center is gradual and is varied to suit the abilities and needs of the individual intern. Attendance at staff case conferences gives the first view of the actual working of the Center. Another important introductory activity in a community psychological service consists of visiting the agencies in the community with which the center works most closely: schools, juvenile court, child placing agency, family agency, welfare department, etc. These visits give two kinds of information: (1) the programs of the agencies and (2) the agencies' view of the Center and its services.

[1]Paper read at a roundtable on the supervision of clinical interns. American Psychological Association, Pennsylvania State College, September, 1950.

From the very beginning it is made clear that psychologists do not function in isolation, but in intimate contact with many aspects of the community. The social workers play a major role in preparing the interns for the visits and in discussing the implications of their visits.

During the first few days an intern is on the job, the supervisory staff can come to decisions regarding types of cases he is ready to handle. Such factors as feelings of confidence in interpersonal relations, adequacy of technical skills, and complexity of the client's problem are considered in the initial assignment of cases. The first cases are given an intake staffing, and the kind of information needed for making further decisions is determined. From the very beginning the intern is encouraged to assist in planning the work that he will do on any particular case. There may be individual planning conferences between intern and senior staff member if needed following the staff conference. There is an informal atmosphere in the "team" work at the Center, and it is made evident that seeking help or consultation is usual among all staff members. The intern does not feel that he is being overdependent in seeking help. After the diagnostic work has been carried out, there is another staff conference at which the findings are presented and treatment or disposition plans discussed. All staff members may take part in this discussion. The intern is responsible for presenting his findings and supporting his conclusions and recommedations. The intern is responsible for writing a report of his psychological study for the case record. When the Center's job is essentially one of developing treatment or disposition plans, the senior staff member (psychologist or psychiatrist) presents these to the referring agency or individual by means of a written report or an interagency case conference. At the end of the year the intern is expected to be able to represent the agency on his own either through report or conference.

The problems encountered by the intern occur in the following areas: technical skills, interprofessional relations, ethical decisions, and personal adjustment. The supervision he receives must help him in all of these areas.

The internship center does assume that the intern will have at his command a variety of examining skills when he arrives. It is possible that specific deficiencies may need to be remedied, but this should be easy to do if he has had an adequate foundation in his earlier clinical training.

The emphasis of the diagnostic work at the Center is on providing an understanding of the individual and his problem as he is referred to us. The most important thing that can be learned in this phase of the internship is the ability to relegate the various psychological tests and scores to their rightful place as aids to the clinician.

The preparation of reports for teachers, social workers, physicians, and public health nurses, conveying the results of psychological study or treatment in a manner that is helpful, is a difficult but important part of a psychologist's task in a community guidance center. It is relatively easy to fall back on our psychological jargon, the use of which may be impressive, but frequently it results in confusion and failure to communicate. It is important, however, to permit the flowering of individual styles of reporting, though not at the expense of clarity and reliability. Too often successful supervision is measured by the degree to which the trainee masters the clinical clichés of his supervisor.

About a third of the Center's time is devoted to some form of treatment—play therapy, group therapy, treatment of parents whose children are in therapy, or psychotherapy of adolescents and adults. The intern is very early assigned cases for treatment, usually children for play therapy.

Although there has been some pressure to institute a more or less formal play therapy seminar early in the intern year, we have postponed this until the last part of the year. We have tried to keep the interns focused on the individual cases and their own problems in handling the actual interpersonal behavior that occurs, rather than permitting a continuance of the intellectual approach. With some interns handling some cases, it is necessary to have a supervisory session following each therapy session for an extended period of time. It has been our experience that, within the supervisory framework we have developed, interns can quickly become aware of their own emotional reactions as therapists and utilize those re-

SUPERVISION OF INTERNS IN A COMMUNITY GUIDANCE CENTER[1]

JOSEPH E. BREWER
WICHITA GUIDANCE CENTER

THE intern training program that I will discuss has developed through the interaction of two forces. One has been the need for trained personnel—a need which was difficult to satisfy because psychologists in training had few opportunities to learn how to function in community psychological work. The Board of Directors of the Wichita Guidance Center realized early that it must devote some of its time and money to training personnel for replacement and expansion. The second force has been the thinking of the APA through its Committee on Graduate Training in Clinical Psychology. This thinking has evolved the general outlines of a professional training program that accords to clinical psychologists a highly responsible service function in collaboration with other professional personnel. The internship period has been structured as occurring in the third year of a four-year doctoral program.

The Wichita Guidance Center internship is designed to form an integral part of a doctoral training progam. It is further designed to give the broadest emphasis possible to this training by experience in a community psychological service, in sufficient extensity and intensity to prepare the trainee for professional competence at a high level of responsibility. We are not interested in developing technicians as such, but are interested in developing technical competence in clinicians. This requires that a full measure of responsibility be given the intern at as early a moment as possible in his year with us. It is important that judgments about the interns' personal integrity and sense of professional responsibility be made early and that assignment of cases and degree of responsibility permitted be gauged accordingly.

The operation of the Center has been described elsewhere. The average monthly caseload is 176. The age range of clients is from two-month infants to adults. The median age level is nine years. The staff to which the interns come consists of three senior psychologists, two junior psychologists, two social workers, two remedial teachers, and two part-time psychiatrists. A staff of this size can handle four interns.

The formal structure of the supervision received by the interns consists of (1) weekly individual conferences with the supervisor of the training program, (2) weekly group meetings with the supervisor, (3) weekly group meetings with each of the psychiatrists, (4) one to three staff case conferences per week, and (5) conferences by appointment with other staff members regarding individual cases. The interns are responsible to a supervisor for their general conduct, and for their program as a whole. Their clinical services are in the beginning under the direct control of a senior staff member who carries case responsibility. Later in their year they may receive only administrative supervision as would any other junior staff member.

The introduction of the intern to the services performed by the Center is gradual and is varied to suit the abilities and needs of the individual intern. Attendance at staff case conferences gives the first view of the actual working of the Center. Another important introductory activity in a community psychological service consists of visiting the agencies in the community with which the center works most closely: schools, juvenile court, child placing agency, family agency, welfare department, etc. These visits give two kinds of information: (1) the programs of the agencies and (2) the agencies' view of the Center and its services.

[1]Paper read at a roundtable on the supervision of clinical interns. American Psychological Association, Pennsylvania State College, September, 1950.

From the very beginning it is made clear that psychologists do not function in isolation, but in intimate contact with many aspects of the community. The social workers play a major role in preparing the interns for the visits and in discussing the implications of their visits.

During the first few days an intern is on the job, the supervisory staff can come to decisions regarding types of cases he is ready to handle. Such factors as feelings of confidence in interpersonal relations, adequacy of technical skills, and complexity of the client's problem are considered in the initial assignment of cases. The first cases are given an intake staffing, and the kind of information needed for making further decisions is determined. From the very beginning the intern is encouraged to assist in planning the work that he will do on any particular case. There may be individual planning conferences between intern and senior staff member if needed following the staff conference. There is an informal atmosphere in the "team" work at the Center, and it is made evident that seeking help or consultation is usual among all staff members. The intern does not feel that he is being overdependent in seeking help. After the diagnostic work has been carried out, there is another staff conference at which the findings are presented and treatment or disposition plans discussed. All staff members may take part in this discussion. The intern is responsible for presenting his findings and supporting his conclusions and recommedations. The intern is responsible for writing a report of his psychological study for the case record. When the Center's job is essentially one of developing treatment or disposition plans, the senior staff member (psychologist or psychiatrist) presents these to the referring agency or individual by means of a written report or an interagency case conference. At the end of the year the intern is expected to be able to represent the agency on his own either through report or conference.

The problems encountered by the intern occur in the following areas: technical skills, interprofessional relations, ethical decisions, and personal adjustment. The supervision he receives must help him in all of these areas.

The internship center does assume that the intern will have at his command a variety of examining skills when he arrives. It is possible that specific deficiencies may need to be remedied, but this should be easy to do if he has had an adequate foundation in his earlier clinical training.

The emphasis of the diagnostic work at the Center is on providing an understanding of the individual and his problem as he is referred to us. The most important thing that can be learned in this phase of the internship is the ability to relegate the various psychological tests and scores to their rightful place as aids to the clinician.

The preparation of reports for teachers, social workers, physicians, and public health nurses, conveying the results of psychological study or treatment in a manner that is helpful, is a difficult but important part of a psychologist's task in a community guidance center. It is relatively easy to fall back on our psychological jargon, the use of which may be impressive, but frequently it results in confusion and failure to communicate. It is important, however, to permit the flowering of individual styles of reporting, though not at the expense of clarity and reliability. Too often successful supervision is measured by the degree to which the trainee masters the clinical clichés of his supervisor.

About a third of the Center's time is devoted to some form of treatment—play therapy, group therapy, treatment of parents whose children are in therapy, or psychotherapy of adolescents and adults. The intern is very early assigned cases for treatment, usually children for play therapy.

Although there has been some pressure to institute a more or less formal play therapy seminar early in the intern year, we have postponed this until the last part of the year. We have tried to keep the interns focused on the individual cases and their own problems in handling the actual interpersonal behavior that occurs, rather than permitting a continuance of the intellectual approach. With some interns handling some cases, it is necessary to have a supervisory session following each therapy session for an extended period of time. It has been our experience that, within the supervisory framework we have developed, interns can quickly become aware of their own emotional reactions as therapists and utilize those re-

actions that contribute to therapeutic progress and modify those that interfere.

Greater caution is needed in planning for treatment work with parents and other adults. The treatment relationship with a parent necessitates making decisions that cannot be placed on any other shoulders. Parental pressure for specific answers regarding a child's behavior tempts the young therapists to be dogmatic and authoritative.

Training in community education and preventive measures has been more difficult to organize and carry out. A community guidance center attempts to use the individual case referred by teacher, social worker, physician, nurse, etc., as a teaching case with those who refer it. Interns at the Center are brought in contact with many community agencies through individual cases with which they work. They are introduced to the community-wide organization of all health and welfare agencies. Some are given experience in appearing before groups to lead discussions on mental health topics. Any intern who has the interest and ability can find opportunities for considerable experience in this field.

The following figures give some idea of the volume of work of an intern: During the first three months of the internship year just ended, each of the four interns at the Center handled from 17 to 24 different cases. They gave from 35 to 59 different psychological tests; they held from 7 to 34 play therapy sessions, from 9 to 22 counseling or parent-interview sessions, and from 0 to 14 remedial-reading sessions. During the next six-month period, the average monthly caseloads for the four interns were 15, 16, 19, and 24. The average numbers of interviews per month were 54, 50, 45, and 52.

Each part of the supervisory process serves a specific function. The staff conferences submit the intern to the same type of supervision that all staff members receive in the group review of clinical activities

The group conference with the supervisor provides a means of meeting the problems faced by all the interns. This group meeting has had a minimum of formal structuring and has been used for discussion of theoretical and philosophical implications of clinical work, as well as for dealing with practical problems faced by the interns. It has been felt that, given an atmosphere of freedom to propose and reject, the embryonic clinician will normally choose wisely and ask for help where he really needs it.

The individual sessions with the supervisor have more or less jokingly been referred to by the interns as their "therapy hour." This may be its use at times by some interns. It is used for review of specific cases, problems of staff relationship, for formulating and evaluating research plans, and for working out problems of relation to university graduate departments.

The intern year should give the student the opportunity to observe ethical behavior and make ethical judgments relative to his own behavior. There are innumerable instances where the confidentiality of information must be maintained and yet a professional opinion given to teachers, social workers, physicians, and nurses. One inhibition that is difficult for the new intern, whose whole energy is given up to his clients, is refraining from discussing his clients with friends outside the staff. It is essential that he have a chance to discuss his new experiences and the interesting sidelights or startling items with members on the staff so that his need for communication will not be unresolved.

Supervision in the etiquette and form of interprofessional relations is carried out at two levels. One is the intrastaff level, where the intern learns to work as colleague with social workers, psychiatrists, and teachers. This involves learning the language of the other professions and their role in the staff organization. Second, interns achieve satisfactory relations with other professional personnel through case conferences at which physicians, social workers, teachers, probation officers, nurses, etc., are present, the mutual interest in the particular client requiring them to understand one another.

Through the supervised experience described here, the clinician in training can achieve the objectives of an intensive and extensive experience with individuals presenting psychological problems. He will have begun the process of developing clinical competence and professional responsibility.

Received November 30, 1950.

SUPERVISION OF CLINICAL PSYCHOLOGY TRAINEES IN A NEUROPSYCHIATRIC HOSPITAL[1,2]

A. L. CARP

VETERANS ADMINISTRATION HOSPITAL, FT. CUSTER, MICHIGAN

IN the emerging field of clinical psychology it has become apparent that many trainees are receiving a large part of their training in neuropsychiatric hospitals and will later find jobs and careers in such institutions. The APA Committee on Training in Clinical Psychology has specified that one year of internship be required for the Ph.D. and that this internship shall be under supervision. Similarly, the Veterans Administration has structured its program to insure a four-year supervised training program. Neither the APA nor the VA has, as yet, prescribed the type of supervision, although the APA has attempted to grade its adequacy on the basis of the experience and background of the supervisors. In this symposium, I assume that we will attempt to discuss some of the problems inherent in supervising clinicial psychologists. I will attempt to limit my remarks to supervision as distinct from training, although I recognize the artificiality of such a distinction.

In developing practica for internship programs, one of the main problems is to provide students with a rich and varied experience, with due regard both for the needs of the particular patient and of the particular student.

As with the medical profession, the clinical psychologist must develop a professional and clinical attitude toward the patient. At all times, training must be subordinate to the therapeutic goals.

In distinguishing between supervision and training, I made the assumption that when the trainee comes to the hospital he has already mastered, at the university, the mechanics of the techniques he will utilize. If he is attempting to serve in a diagnostic capacity, he must have learned the administration, scoring, and the rudiments of interpretation of his diagnostic instruments prior to applying them to patients. This is, of course, the responsibility of the training institution or university. He has learned these techniques by taking them himself, and administering them to friends and to other nondisturbed individuals. He has had sufficient practice so that he can pay minimal attention to the testing itself and maximal attention to the responding patient.

In his role as a diagnostician at the hospital the trainee has two main responsibilities: (1) He must get an adequate test to provide the basis for later interpretation, and (2) he must conduct the testing in such a manner that the testing experience is, at the least, not harmful and possibly of benefit to the patient. In my experience, the second aspect has been most neglected, and students are most deficient in it. In this area the most careful supervision is necessary. The supervisor must insure that the trainee functions not only as a tester, but also as a clinician, that before the trainee starts testing he is aware of the patients's attitude toward the testing situation, and can make some guesses as to what effect it might have. He should be aware of the more common kinds of resistance. His clinical orientation should determine the oft-asked question, "How much time should be spent with the patient before beginning the test?" Similarly, the student must be aware of the ending of the testing situation, so that he does not pick up his material and disappear, but makes an attempt to reduce any anxiety or hostility engendered.

[1]Sponsored by the VA and published with the approval of the Chief Medical Director. The statements and conclusions published by the author are a result of his own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

[2]Paper read at a roundtable on the supervision of clinical interns. American Psychological Association, Pennsylvania State College, September, 1950.

Diagnostic testing is usually the student's first approach to actual clinical functioning, and its supervision is important because it introduces the clinical approach early in the student's training. "Patient-oriented" supervision and indoctrination may help to mitigate the criticism sometimes applied to clinical psychologists, that they are not overly concerned with the patient, but merely interested in getting an accurate score. The student's training in the academic setting ordinarily does require that he learn a standard, objective, precise manner of administering the test. He must not lose this objectivity, but he must learn to temper it with a feeling for what the test means and does to the patient. In a neuropsychiatric hospital, much of the diagnostic testing is done with psychotic individuals who, on the surface, may appear to be grossly uninterested in the test or in the examiner. Still, the testing procedure may have a pronounced effect on such patients. More than one clinical psychologist has become incorporated into a paranoid patient's delusional system, because of his failure to appreciate the clinical significance of the testing situation. Another aspect of the supervisor's responsibility in the area of diagnosis is equally important. The supervisor should be responsible for the interpretive report written by the trainee, not only for the accuracy of the interpretation, but also for the translation of psychological jargon into language that will be meaningful for those who will make use of it. In summary, the supervisor of diagnostic practice has a responsibility for insuring not only that the test is properly administered and competently interpreted, but also for making the student feel at all points the significance of the testing procedure for the patient.

In the area of therapy, even more than in the diagnosis, we have a responsibility to provide therapeutic experience for the trainee, and at the same time to insure that the patient does not suffer in the process. There is also the problem of insuring that the experience is not too traumatic for the student. Here lies one of the main responsibilities of the supervisor.

The type of supervision and the methods employed will depend on one's philosophy of training. Supervision may consist mainly of a therapeutic experience for the trainee, or it may be directed more toward the mechanics and problems of therapy. The many different orientations toward therapy that supervisors as well as trainees bring to the situation increase the complexity of the problem.

Without attempting to solve all of these difficult issues, I would like to describe the program as it is being carried out at one installation. We conceive of therapy training as proceeding through the entire four years of the clinical program. In the first year, the students are primarily occupied in taking admission histories from newly admitted patients. It is the supervisor's responsibility to insure that, during this experience, the student becomes aware of the patient's reactions to the different questions and areas of the history. He must also become explicitly aware of the interpersonal relationship between himself and the patient. He must sense his reactions and feelings toward the patient, the patient's reactions and feelings toward him, and the meaning of the history-taking to the patient. We feel that in this area the trainee can be introduced to the interpersonal situation with disturbed individuals with minimal harm to the patient or to himself, and with maximal opportunity for learning.

In the second year of the clinical program, the student operates primarily in the diagnostic field. Here the supervisor, in addition to his other functions, again makes the student aware of the interpersonal relationship operating in the testing situation. Also in the second year, each student attempts to establish a relation with a chronically hospitalized patient. The student's goal is to know the patient and to establish a free relationship with him. He is not expected to do psychotherapeutic interviews nor to cure the patient. Here again, material is provided for discussing both the trainee's reactions and the patient's reactions.

In the third and fourth years, the trainee devotes more time to psychotherapy with less chronic patients and with patients who are expected to improve because of their therapeutic sessions. In that phase, the supervisory relationship will depend more on the particular

philosophy of the supervisor, but with his gradual introduction to observing both his own and the patient's reactions, the student should be able to profit from any type of adequate supervision.

One question deserves to be raised even though it cannot be answered. What are the qualifications of the good supervisor? There is little research in this area, although there is some agreement that the qualifications of a therapist may be different from those of a diagnostician or of a researcher. Research in this area is urgently needed to replace the present policy of availability and degree status as the necessary qualifications of a supervisor.

Received November 30, 1950.

TRAINING IN PSYCHOTHERAPY[1]

O. H. MOWRER
UNIVERSITY OF ILLINOIS

RESPECT for the privacy of a man's own thoughts and tolerance of personal idiosyncrasies are among the basic values of our culture. So long as a person "behaves himself" within rather broad limits and "does his work" effectively, we grant him full privileges as a citizen and sanctity within the confines of his feelings, desires, and opinions. This, in a word, is our conception of personal freedom.

If, however, a person comes to us for help in connection with a personality disorder, the situation is altered. We allow, indeed, encourage, the person to reveal his innermost self. But our graduate students in clinical psychology come to us, not for treatment, but for *training*; and it is not as yet entirely clear what status we should accord them and what experiences we should try to provide for them.

Psychoanalytic therapists have taken a position at one extreme of a continuum, holding that treatment and training are not fundamentally different; hence the tradition of the didactic analysis. At the opposite extreme is the position that psychotherapy is something that can be done by anyone on the basis of a few simple rules, and without regard to the personal qualifications of the practitioner.

It is unlikely that the latter position will long prove defensible, but neither is it certain that intensive didactic therapy is essential in all instances. Rather than stress any fixed, invariant training routine, what we need is a trustworthy method of assaying something for which we have no better term than *personal maturity*. This alone is not enough; there must also be an active interest in the process and goals of therapy, and there must be conceptual sophistication. But without stability and strength of character, these other attributes come to naught.

This is not to say that anyone who has experienced developmental problems is necessarily eliminated from the practice of psychotherapy. A person who has had a neurosis, perhaps quite a severe one, and who has gotten over it, and who knows *how* and *why* he got over it, may be the salt of the earth as a therapist. But certainly anyone who currently suffers from neurotic manifestations is unwarranted—regardless of his diplomas or formal training— to practice psychotherapy.

Our problem, then, is how to select for training in this field persons with the requisite motivation and maturity, actual or potential. Already we can select with reasonable adequacy for sheer intellectual ability; and by ordinary instructional methods we can insure satisfactory familiarity with fact and theory. But proficiency as a therapist calls for talents which at present lie beyond the scope of established tests or formal teaching methods.

One reason for stressing motivation in connection with the practice of psychotherapy is clear: therapy is not something that one can do routinely, without sustained attention and thought. It is a thoroughly creative enterprise and calls for the best one has by way of concentration and intellectual prowess. But further consideration of the problem reveals a dilemma: namely, how to achieve active interest and yet avoid personal involvement. Dr. I. N. Korner [1] has recently discussed this problem in a highly competent and informed manner; and his analysis and recommendations need not be repeated here.

Perhaps more needed at present is consideration of the problem of personal maturity and how it can be properly assessed and fos-

[1]Paper read at a roundtable on the supervision of clinical interns. American Psychological Association, Pennsylvania State College, September, 1950.

tered in the would-be therapist. Among just-published books is one engagingly entitled: *How To Be Happy Though Young.* Also needed is counsel on: How to be a competent psychotherapist though young. While aging does not always mean greater personal maturity, the correlation is well above zero; and, other things equal, a man or woman in the early thirties is a better candidate for training in this field than is a person a decade younger. At the present time, we commonly find students of twenty-two or twenty-three in our therapy classes and the associated practica. In comparison with what can be done by way of training a man or woman of thirty or thirty-five with equally good endowment, my experience with these younger students has not been very encouraging.

I once asked the dean of a theological seminary what he did with all his bright young graduates while waiting for them to develop enough maturity and wisdom to be effective ministers. The somewhat cynical reply was that this was no problem at all; the middle-aged and elderly women in a congregation are always delighted to have a boyish-looking man in the pulpit whom they can "mother." The question, said the dean, is what to do with ministers when they begin to get gray hair and have a few wrinkles.

But in the field of psychotherapy, youth is a real problem. Here, perhaps, is another argument to add to those already advanced (at the Boulder Conference [2] and elsewhere) for feeling that intensive training in therapy ought to be postponed until the postdoctoral level. Another possibility which has been suggested is that we start our therapy students with children. Because of the age differential in such circumstances, most students do, indeed, feel more comfortable and do better work with children, but another complication quickly arises. We know that neurotic children nearly always have neurotic parents; and for every patient we have in play therapy, we are therefore likely to have one or more persons in adult therapy. An ideal plan is one in which the student carries the child in therapy and the supervisor works with one (or both) of the parents. This provides an extraordinarily valuable learning experience for the student, but it is evident how demanding this plan is of the time of supervisors.

Thus far we have been rather circumstantial and pragmatic in our emphasis upon the personal maturity of the therapist. More specifically, what does this imply? The problem, as I see it, breaks down into two parts: one having to do with *interpretation* and the other with the *transference.*

At the outset of an orthodox Freudian analysis, the patient is told to "free-associate." In the beginning this is the one thing that the patient cannot and will not do. This is an objective to work toward as therapy progresses, but in the beginning it is the *therapist* who must do the free-associating. The central strategy of neurosis is dissociation, and the essence of an interpretation is that it reverses the dissociative trend; it replaces dissociations with associations, disconnection with connections, unconsciousness with awareness insight. And it goes without saying that a person with serious blocks and barriers in his own personality does not have the requisite mental flexibility to do effective therapy. Suppose the patient is a nineteen-year-old college boy who is suffering from masturbation guilt and insecure masculine identification. Suppose that the would-be therapist is a twenty-three-year-old unmarried male graduate student. The chances are that he, too, will not have fully resolved the masturbation conflict and that he, too, will not be disposed to speak or even think about matters which *must* be thought about and discussed if therapy is to prosper.

Or, let us imagine a youthful student-therapist of either sex trying to work with a highly intelligent, verbally aggressive neurotic patient of forty. Assume that sufficiently penetrating interpretations have been made to arouse lively resistance and stimulate transference strategies. Unless he is a very remarkable young person indeed, there is a high probability that the student's own defensives will be activated by the patient and that he will be pitched headlong into fatal countertransference behavior with the patient. Radical therapy is never, I believe, accomplished without the emergence of transference reactions; and it is not evident to me how the onslaught of these strategies can be withstood, to say

nothing of being turned to therapeutic account, unless the therapist has not only the necessary technical understanding but also a considerable store of inner assurance, security, and conviction. I do not say that a graduate student in his early twenties cannot have or acquire the purely personal qualifications which are prerequisite for therapeutic proficiency; but the odds are against him, and we would perhaps do well to devise a system of training in which such students are not prematurely exposed to experiences which, in another five years, they might successfully manage.[2]

Every experienced therapist expects to lose contact with a certain number of patients, perhaps one in four or five. Successful therapy depends not alone upon the competence of the therapist; there must also be at least a minimal degree of therapy-readiness on the part of the patient. If, therefore, a student therapist loses his first, or perhaps also his second patient, he can rationalize such an experience; but if he, himself, is not to be lost as a psychotherapist, we must make sure that he does not have too many negative experiences with patients. Anything which will promote a higher incidence of successes for the beginning therapist deserves our careful consideration.

Close supervision of the beginner has been found to be generally desirable. As an ideal, it is not too much to say that the neophyte therapist should spend two hours with his supervisor for every hour he spends with a patient. A didactic therapy tends to reduce the amount of supervision which a student will require, but whether it will save enough in supervision to make up for the time which the didactic therapy itself requires is an open question; certainly the answer will vary in individual cases. Recordings or, somewhat better still, sound movies made by another therapist are extremely useful to the student in advance of his being assigned his first case. By these means he can have the experience of being *in* the therapeutic setting without, however, having full responsibility for the situation. And when the student starts work with cases of his own, recordings are invaluable both to the student and to the supervisor.

In this brief presentation we shall have no time to discuss other issues which arise in connection with training in psychotherapy. Nor shall we be able to consider any of the many pressing questions that beset us at the level of content and theory. In the moment that remains, may I therefore summarize the central point which I have tried to make in this paper. Therapy, it appears, is not an enterprise which can be conducted on the basis of a few routine rules or a single, fixed formula. It calls for at least as much resourcefulness and versatility as, let us say, successful parenthood, good teaching, or other professional tasks. Moreover, in therapy our goals are not purely negative. Contrary to a common impression, the objective is not primarily that of undoing inhibitions, releasing pent-up biological impulses, extinguishing bogus fears, or softening up the superego. Therapy is preeminently an integrative, synthesizing, ethical task which at times is definitely upgrade work against inertia and resistance; and it calls for the best that we can bring to it in terms of self-awareness, technical sophistication, pro-

[2]Dr. Abraham Carp has called the author's attention to a special circumstance prevailing in mental hospitals. Here there is such a great distance between the status of a patient and even the most youthful staff members that some of the problems encountered in outpatient therapy may not necessarily arise. Here, as Dr. Carp observes, a crucial question is, "Who can go out through the gate at night?" Or, "Who has the keys?" Expressed a little differently, the mere fact of confinement does something to counteract the age variable and to create a status discrepancy between patient and therapist which might otherwise not exist. I am sure this is an important consideration, but one wonders if it may not be offset somewhat by the greater inherent difficulty in doing therapy with psychotics than with the types of persons ordinarily encountered in outpatient practice. In other words, it does not seem likely that the rather artificial status or "maturity" which the mental hospital setting automatically confers upon the student therapist necessarily meets the fundamental problem with which we are here concerned. A related issue has been raised by a recent symposium on the question of medical vs. "lay" psychotherapy. Here one of the medical participants put great emphasis upon the advantage which a medical therapist, however youthful, enjoys solely by virtue of the fact that he carries the venerated and respect-commanding title of "doctor." This would be an interesting issue to investigate empirically, but one can conjecture that in the end even the medical therapist has to win the battle of the transference, not on the basis of some ascribed role, but on the basis of demonstrated technical competence and inherent qualities as a human being.

fessional dedication, and personal integrity. We must not underestimate the difficulty of wise selection and the time-consuming, costly nature of training of competent personnel in this field. Both in terms of the age factor and the time requirements, it is doubtful if we shall be fully satisfied with the results of our efforts along these lines until we have instituted either postdoctoral therapy training programs or have established advanced internships which offer special facilities in this connection.

Received November 30, 1950.

REFERENCES

1. KORNER, I. N. Ego involvement and the process of disengagement. *J. consult. Psychol.*, 1950, **14**, 206–209.
2. RAIMY, V. C. (Ed.) *Training in clinical psychology*. New York: Prentice-Hall, 1950.

THE PSYCHOLOGIST'S BEHAVIOR AS AN AREA OF RESEARCH[1]

SEYMOUR B. SARASON
YALE UNIVERSITY

IT will be assumed in the following discussion that there is agreement that two of the most important objectives in a clinical training program are to get the student (*a*) to be aware of how his own reaction tendencies are influenced by the kinds of clinical situations and patients he sees and (*b*) to learn to deduce covert behavior from that which is directly observable. Before taking up the implications of these objectives for the nature of training procedures, it may be advisable to take note of some obvious characteristics of the beginning clinical student. He is in his early twenties, has had little or no contact with clinical material, has no more insight into his own behavior than other college graduates, and his conception of the objectives of clinical training probably is in part unrealistic—to many, such training consists of learning procedures which enable one to find out why *other* people behave as they do. It is necessary to mention these characteristics because they emphasize at least two things: (*a*) the student has much to learn and unlearn, and (*b*) the attainment of the objectives listed above will take much time and will require experiences in situations which are conducive to the acquiring of new reaction patterns.

To the above banalities one might also add that the student requires the supervision of an experienced person. But how much agreement is there about the nature and extent of supervision? Do our present supervisory practices achieve our purposes? What does the student learn by going over a case in the office of his supervisor? What does he learn by presenting his case at a staff conference? What does he learn in those training situations where some of his cases are "checked" by and discussed with other students? What does he learn when there is no supervision? Probably —and this a statement of faith—the student gains something from each of these types of learning situations. But does he learn that which would be considered relevant to the two important objectives listed in the beginning of this paper? The most acceptable and discouraging answer would be that we don't know. The writer's opinion is that the student gains little or nothing.

From current supervisory practices the student is not in a position to learn how his observations and written statements about a patient are in the nature of selective reporting. He has no opportunity to learn what factors unique to him are responsible for his selective perceptions. He is not in the position to learn how to distinguish between conclusions based on observation and deduction. In short, the student cannot sufficiently become aware of the relation between his verbalized and written conclusions and his own behavior. How many students learn to examine their own reactions? Why does the student like this patient and dislike that one? Why should he have this phantasy about this patient? Why should he experience apprehension and insecurity with a particular patient? Why is he bored with this patient? Why does he think the patient is anxious or evasive or hostile or unresponsive? The purpose of asking such questions is not self-criticalness for its own sake but for the sake of relating the psychologist's overt and covert behavior to the stimulus situation. What external cues are responsible for the psychologist's reaction? To what extent is the reaction a function of the psy-

[1]Paper read at a roundtable on the supervision of clinical interns. American Psychological Association, Pennsylvania State College, September, 1950.

chologist's private and unique experiences? To what extent would other psychologists react similarly? In his function as a diagnostician, how much of the clinician's conclusions derive from his formal test procedures and how much from deduction from overt behavior? To those who would maintain that one can make many more reliable and valid conclusions from formal test responses, one might answer that if we spent as much time calibrating the psychologist as we do the tests he uses, the accuracy and extent of our predictions might be considerably enhanced. If a therapist went only by what the patient *said,* he would get nowhere. How does the therapist go from the verbalized to the nonverbalized attitudes and needs of the patient? How does he learn to do it? That some people are born with "the gift of insight"—that they are a breed apart—is a gratuitous assumption in the light of our ignorance in the matter. Such an assumption at this time indicates how little we know about the psychology of the psychologist.

It might be thought by some that many of these problems would be obviated if each student could be psychoanalyzed. While such an experience would undoubtedly be of help, for several reasons it cannot be considered a panacea. First, not everyone benefits to the same degree and in the same way from an analysis. Secondly, the new learning which occurs during analysis may largely take place in relation to people and situations which are rather different from the people and situations the student will encounter professionally. Thirdly, if what the individual has learned in analysis is not practiced and reinforced in his professional work, then the amount of transfer of training will be small. In other words, the conditions of learning to become a clinical psychologist must be such as to maximize the effects of learning in analysis. The two situations supplement each other; they cannot be substituted for each other. This question is probably an academic one because there is little likelihood that in the foreseeable future, analysis will be made available to any but a few students. The attainment of the objectives discussed in this paper is primarily the obligation of the supervisors.

In the most recent report of the APA Committee on Training in Clinical Psychology, on recommended standards for practicum training [2], there is a listing of the teaching procedures "which seem to contribute most to the total development of the practicum student." The first of these methods follows:

> Controlled or recorded observations of the behavior of patients. These observations should be carried on in as many different settings as possible. In the mental hospital, for example, they may be done in the wards, day-rooms, grounds, testing and therapy rooms. Such observations should be scheduled early in the practicum period. Sharp, precise observations are prerequisite to meaningful interpretations. In addition, the student should be asked to interpret the significance of the behavior observed and to discuss his observations with the supervising psychologist or the professional person in the best position to criticize and evaluate the student's work [2, p. 601].

This quotation seems to contain an ambiguity: Are the situations which the student observes discussed with a supervisor who also has observed the same situation? If this is not what is meant—if this is not mandatory—then it is difficult to see how discussing observations with a supervisor who was not in the situation helps the student to become a precise observer.

The next three methods taken up by the Committee concern supervision of report-writing, the use of case conferences, and the role of staff meetings. While these are important, they do not appear to be relevant to the objectives being discussed here. The fifth teaching method follows:

> The frequent observation by the supervisor of the student's actual work with patients, followed by criticism, is desirable. Misunderstandings and errors are often corrected only in this way. The use of one-way vision screens, sound recordings and similar methods is helpful in this respect [2, p. 602].

Two questions might be asked about Committee's statement: Why is frequent observation only *desirable?* Is the correction of errors and misunderstanding the only or the most important result of direct observation of the student?

The sixth method discussed by the Commitee concerns the teaching of psychotherapy. Concerning this area the report states:

> The use of recorded and transcribed psychotherapy sessions and the use of one-way vision screens and inter-communication systems are of great technical aid in teaching psychotherapy but they can only supplement individually supervised practice [2, p. 602].

This writer feels that several comments might be made about the section on methods in the Committee's report. First, the necessity of frequent observation is not sufficiently emphasized—it is the difference between saying that something is desirable or that it is essential. Secondly, the functions or uses of the method of direct observation either seem to be narrowly conceived or not sufficiently spelled out.

It should be made clear that the important problem is not how to train the student to be a precise observer in the sense that he will come from a situation with an inventory of all that has happened. The important problem can be put in the form of questions: How can we get the student to become aware of the selective nature of his observations? How can the student learn to distinguish between overt behavior and deductively obtained statements of covert behavior? How can we help the student to become aware of what constitutes relevant proof or disproof of a deduction and of the necessity for a deliberate consideration of alternative explanations? How can we train the student to use his own reactions to a situation as a means of understanding the reactions of the patient?

We know very little about the answers to these questions. This state of affairs is a reflection of psychology's neglect of the psychologist's behavior as an area of research. If one assumes either that the psychologist's mental processes do not particularly influence the conclusions he draws or that one will gain relatively little in attempts to sharpen his observations and the validity of deductions drawn therefrom, then the problems discussed here are unimportant. Those who make such or similar assumptions will probably continue to view the goal of clinical training primarily as the acquiring of a body of external knowledge and procedures which is independent of how and by whom they are used.

Our ability to predict behavior is rather poor. In attempting to ascertain the reasons for this sad state of affairs, the writer has concluded that in predicting from one situation to another, we have neglected some important determinants of behavior. Only very infrequently have researchers taken into account those determinants which are not necessarily discernible either from responses to formal test items or from the subject's articulated verbalizations, but rather are deducible from overt behavior: attitudes toward self, others, situation, etc.[2] If one assumes that such attitudinal factors are unimportant, then there is no problem. If they are assumed to be important, how does the psychologist make relevant conclusions about them? What is the nature of the deductive process whereby one goes from overt to covert behavior? The clinician makes such deductions, but for some reason or other such data are not considered in whatever actuarial procedures are employed in prediction studies. When one sees how the data which the trained observer can contribute are neglected—data which are highly relevant to the problem of prediction, since they help define the nature of the situation and enable one to determine to what other situations one is entitled to predict—one can only reiterate the need for more attention to be given to the observer as a rich source of valuable data.

Received November 30, 1950.

References

1. Sarason, S. B. The test-situation and the problem of prediction. *J. clin. Psychol.*, 1950, **6**, 387–392.
2. Standards for practicum training in clinical psychology: tentative recommendations. APA Committee on Training in Clinical Psychology. *Amer. Psychologist*, 1950, **5**, 594–609.

[2]This point is discussed more fully in another paper by the author [1].

OSCILLATION AND PROGRESS IN CLINICAL PSYCHOLOGY[1]

MAUD A. MERRILL
STANFORD UNIVERSITY

CLINICAL psychology began in a formal way with a name and place of residence when Lightner Witmer established at the University of Pennsylvania in 1896 his clinic for the diagnosis and treatment of children who were mental or moral deviates from the normal. It became articulate through the medium of its first journal, *The Psychological Clinic,* which Witmer edited throughout the 28 years of its publication from 1907 to 1935.

This first clinic and the journal which presently became its voice crystalized into institutional reality certain trends current in the psychological developments of that period. Motivational and methodological factors which characterized its early form and functioning have served both to facilitate and inhibit the development of clinical psychology as a profession. It is our present undertaking to see how some of the issues now looming large as crucial in the determination of our future as a profession, came to be part of our thinking.

Clinical psychology came into being in response to the need for service. Its purpose was to make use of the available tools of the new science of psychology to serve human needs—especially the needs of defective and of delinquent children. Such service orientation may in some respects stimulate scientific thinking but in other respects may inhibit it. The demands of service are insatiable. Our problem in 1951 is not essentially different (only more insistent because there are more of us and the demands are immeasureably greater) from the problems we faced at the beginning of our professional lives, the perennial problem of how to promote professional growth while avoiding the stunting effects of too many and too insistent service demands. How are we to keep from being drained of our vitality by too much service and too little basic thinking?

There were new tools in the 1900's for psychologists to use in assessing human abilities—these were the tests that under the exigencies of practical considerations had been devised by Binet and his collaborators. Then it was the measurement of intelligence that came to be thought of as the most important service function of the clinician. How too much reliance came to be placed on the new tools and too much to be expected of a technique which had proved useful for its original purpose is part of the history of clinical psychology and a direct outcome of the always too insistent demands of service—demands which outrun our knowledge about both the causes of variation in human behavior and the control of these causes.

There are new tools in the 1950's, particularly in the field of projective techniques. Already the service demands upon these, too, are far in excess of our factual knowledge about their uses and limitations. Like the older tools, they were devised as diagnostic aids, not to further the scientific understanding of abnormal personality. The pendulum it seems, has swung to the other extreme; whereas Binet had been caught by a Zeitgeist that demanded quantitative measurement in all the sciences, the wave of projective personality-testing has arisen as part of a very vocal protest against quantification. The attempts to measure personality in the same way that had earlier proved so satisfactory for testing intelligence had yielded little of demonstrated value for the clinician, yet the need for techniques for assessing other than intellectual aspects of

[1]Address of the retiring President of the Western Psychological Association.

personality became increasingly urgent as the service demands increased. This unprecedented demand for the services of clinical psychologists has far exceeded the availability of trained personnel with adequate professional skills. The alarmist stage was reached just following the temporary cessation of hostilities after World War II. That fact is abundantly evident in the articles and editorials appearing in the *American Psychologist* and other psychological journals in 1946 and 1947. We were admonished then not to underestimate the gravity of the situation, and it was pointed out that, because of our inability to meet the great public demand for the services of psychologists, the future of clinical psychology was threatened by the likelihood—nay, certainty—that its techniques would be used by "those upon whom we have not set our stamp of approval." "We shall bear the onus of the misuse of these techniques in the eyes of the public," the critics added [7, p. 180]. Binet in the early 1900's and Terman in the 1920's pointed to this same danger in the misuse of intelligence tests by inadequately trained persons—"hordes of untrained workers are likely to use [the tests] for all kinds of purposes that have nothing to do with science" [18, p. 108]. To these alarms sober reflection about current trends adds another still greater, the danger that under the pressure of these exhorbitant service demands we may relax our efforts to evaluate our professional tools, be they test techniques or therapeutic skills, and to assess the validity of our assumptions about the nature of personality.

It was impressively evident in August, 1949, at the Boulder Conference on Training in Clinical Psychology, that the validation of its own techniques and theories was considered to be the most urgent problem now confronting clinical psychology. One of the members of the conference summed the matter up: "What emerged from the discussions was a sober recognition of the fact that, though society may need the services of clinical psychology, what clinical psychology needs most urgently is more dependable knowledge and more valid methods with which to do its job" [15, p. 102]. The conference report, constituting currently the voice of the profession, emphasizes the need for research to the end that we may know just what we are doing when we use the various techniques that clinical psychologists habitually employ. The pressure of day-by-day service demands and our own needs to serve have impelled us to make use of practical devices whose validity has still to be demonstrated. "We have perforce," as the conference report points out a little defensively, "used tools in need of sharpening, and relied on intuition in lieu of data. We have accepted assumptions, sometimes uncritically, sometimes skeptically, and sometimes for the time being only or as 'best available knowledge' " [16, p. 26]. So here we are in the anomalous position of having affirmed soberly on Monday that we must strive for validation and predictability and then, still no wiser on Tuesday, being compelled to accept Monday's doubts as Tuesday's truth. It is usual to think that those who assess sin adequately on Sunday have only to stop sinning on Monday, whereas they probably have to continue in sin for weeks while they find the way out, being involved in contractual obligations and ethical necessities—for good ethics may require sinning when the choice is between two mutually incompatible sins. You just embrace the better sin. This for the clinician is an old dilemma.

So it happened that the growth of clinical psychology was always both facilitated and hindered by its role—it was stimulated by social demands to develop techniques adequate to serve these demands and at the same time progress was restrained by the practical need for habitual and standardized procedures.

Progress in clinical psychology is at best difficult to evaluate because the events which make up the history of its development are too recent to be viewed in the perspective of temporal distance. The leaders of clinical psychology are still our professional colleagues, and many of the chief principles and theories are still lively controversial issues or at most unrepressed recent memories. Still professionally self-conscious and feeling not quite sure of ourselves because, as Woodworth put it, like the psychiatrists we are "in possession of only a modicum of relevant knowledge," we tend to "over-compensate for this feeling."

The prestige of certain practices confers a spurious kind of validity upon them, and we resort to much "verbal staking out of claims" in the therapeutic no-man's-land between psychology and psychiatry [21, p. 4].

Insight into the developmental trends of clinical psychology *in statu nascendi* may be gained by applying here certain criteria of scientific progress [4, 5].

FACILITATORS AND INHIBITORS OF PROGRESS IN CLINICAL PSYCHOLOGY

Clinical psychology has advanced because it has found new tools and techniques. Progress has been fostered by theories, the thinking upon which clinical practices are based, and nurtured by the psychological climate of thought in which its development has taken place. Periods of rapid development, occasioned by emergency demands, have been followed by periods of little apparent gain, the plateaus in our practice curves, when previous gains are being consolidated. Such a plateau in the 20's followed the initial impetus given to testing when these first tools of the clinician were being subjected to evaluation, and the exaggerated expectations of over-enthusiastic users were being reduced in the crucibles of research. This era of the twenties, characterized by Louttit as "the dark ages of mental testing," [12, p. 366] can, on the other hand, be seen as a period during which needed research contributed to a better understanding of the uses and limitations of tests. The tests could be tested by the consistency of their predictions. For instance, psychologists began to determine the extent of agreement between earlier and later tests of the same individuals, and the extent to which other variables affect the obtained correlations. Thus "the dogma of IQ constancy" was reduced to the sounder status of a probability statement. In terms of the most probable eventuality, the IQ is a more valuable tool for the clinician's diagnostic kit than it was in the days when the needs of the believer invested it with a kind of infallibility. This faith was due in part to the fortuitous circumstance that the IQ came into general use only shortly after the rediscovery of Mendel led to the efforts to apply Mendelian principles to the inheritance of intelligence. If you think of "intelligence" biologically as a unit character, measurable pyschologically by tests, and diagnostic significance established in quantitative terms, it is easy to see how a belief in its immutable character came to be part of this concept. Important, too, to remember is how much clinical psychology needed tools for its professional advancement, and even in the search for truth necessity can be the mother of invention.

That the development of clinical psychology has been fostered and also limited by the climate of thought of the time, what Boring has called being in "phase with the times," [5, p. 345] is indicated by the well-known historical circumstances that led to the invention of tests and their promotion to become the much-used and much-abused tools of the clinical psychologist. (In 1939 Louttit still felt that it was necessary in defining clinical psychology to say that it was not psychometrics, that it was not medical psychology, and that it did not deal primarily with the subnormal or abnormal [12, p. 370]). Galton's invention of the mental test, which differed in its brevity and its simplicity from the elaborate psychophysical methods of the experimental psychologists of the period, had been in the interest of assessing human capacities in order to disclose individual differences. Galton had opened a laboratory for inventorying abilities in 1882. Cattell, in the Galton tradition of using tests of the simpler sensory-motor functions, undertook in his new laboratory at the University of Pennsylvania to develop similar tests in 1890. Binet, an experimentalist in the French tradition, brought psychology out of the laboratory when he developed tests for practical mental classification of retarded school children, the extreme deviates intellectually of the Gaussian continuum of abilities which Galton had been at pains to demonstrate. Tests were the tools of the new differential psychology, and differential psychology was interested in the individual deviate. Practical social considerations demanded that something be done in regard to the people who were social misfits. Clinical psychology undertook this task.

That the emphasis came to be placed on intelligence testing was in large part determined

by the availability of particular tools first used in the interests of individual psychology.

THE FIRST CLINICAL ACTIVITIES

Cattell had been the vigorous promoter of differential psychology in America. Both Cattell and Witmer, his successor at the University of Pennsylvania, had been trained in experimental psychology in Wundt's laboratory at Leipzig. Each became concerned, however, not with description of the generalized human mind, but with the application of the laboratory techniques to the assessment of personal capacities in the adjustment of the individual to his environment. In Witmer's hands the psychology of human capacity began to acquire its professional bent. So Witmer's clinic was service-oriented and service-determined. Dedicated to "the restoration of the retardate and the deviate to normality," it espoused a cause. Its journal announced, "*The Psychological Clinic* is propagandist in spirit and welcomes every effective contribution to the cause which it supports." It proclaimed that "the prevention and cure of mental deficiency and defect . . . must be made to hold the center of attention" [20, p. 2], and some of you will remember that it was in the hope of curing mental deficiency that institutions for defectives were first established in this country when Seguin was brought here in 1848 to inaugurate his educational methods at the newly established Massachusetts School for the Feebleminded. Fifty years later it was to become front-page news that feeblemindedness was being cured by special techniques developed in the salubrious educational climate of the Middle West.

Louttit laments the fact that the Witmer clinic and "the things Witmer stood for were far from being the important elements in activities going on under the name of clinical psychology" [12, p. 366]. Nevertheless I think it is clear in the perspective of 1951 that the persistent, if sometimes misguided, efforts to evaluate our test tools, efforts which constituted the bulk of research activities of Louttit's "dark ages," have yielded a better understanding of the uses and limitations of these techniques and have led to a sounder basic theory of intelligence than would have been the case had the clinical practice of the period constituted the chief concern of psychologists interested in problems of individual differences.

The activities that went on under the name of clinical psychology were described in considerable detail in articles which appeared in the early numbers of the *Psychological Clinic* and also in a book, *The Conservation of the Child,* by Arthur Holmes of the Witmer clinic. *The Conservation of the Child* became in the years following its publication in 1912 the guide to Counseling and Psychotherapy for the young clinician. This clinic of Witmer's was the prototype of psychological clinics in the first decade of their existence.

The clinic's central function was mental diagnosis, which was defined as "the classification of the exceptional child according to his mental capacity" [11, p. 52]. Having determined the mental capacity of its child clients, the clinic next proceeded to administer to parents, teachers, and guardians "advice which they could follow out in the home or at such other places as might be suitable." Besides service, the clinic's training and research functions were emphasized. The training program for "psychoclinicists" was to fit them to become examiners in psychological clinics which functioned as "educational dispensaries." The clinic's research function was described as "collecting a lot of data for a lot of children and filing for some future reference" [11]. The published results of "research" so conceived and so executed are poignant reminders that R. A. Fisher's *Design of Experiments* was indeed at that time still twenty-three years away.

The study and treatment of moral as well as mental deviation was of primary concern to the first clinic. Here, too, mental taxonomy was expected to offer the data upon which to base the formula for treatment. Thus, what to do about delinquency depended upon the determination in the individual case of whether or not the delinquency was curable. All moral deviates, according to this view, are either corrigible or incorrigible, and each child client who suffers from moral deviation must be assigned to one or the other category before his disposal, future training, and treatment can be prescribed. Holmes observed that "un-

less the clinician can classify the bad girl or boy presented to him, and upon the basis of the classification advise the best procedure in each instance, he can offer no good reason for his existence" [11, p. 274]. That, consistent with these conceptions of their functions, so many clinicians proceeded to offer good reasons for their existence has formed the chief burden of complaint against the era of mental testing in clinical psychology.

CHANGING VIEWS: DELINQUENCY AND PREDICTION

It is fortunate for clinical psychology that other influences came to bear on its development. Healy's clinic, established in 1909 to serve the Chicago Juvenile Court, initiated an effort to take account of the dynamics of behavior. Three years after the publication of *The Conservation of the Child,* appeared Healy's *Individual Delinquent,* a book which presented a very different and not then widely accepted view of the moral deviate. Many years were to elapse before sound empirical principles were to be established along the lines proposed for the study of child offenders at the Chicago Clinic. It was nearly 40 years ago that Healy described the purpose of his research on child offenders in the following way: "The prime motive for our research into beginnings and causative factors we have ever felt to be the establishment of scientific laws of predictability upon which all sorts of treatment could be rationally planned" [10, p. 16]. I am not sure that Healy would agree that the predictive devices offered by the Gluecks in 1950 in their most recent attempt at *Unraveling Juvenile Delinquency* have established such scientific laws. The Gluecks, using the empirical data which have occurred in the case of delinquents to an extent significantly greater than in the nondelinquents, calculate the chances that individual X will become a delinquent by taking account of the particular combination of factors found in his case. They are making an actuarial prediction, whereas Healy's basic philosophy favors a clinical approach to prediction.

Changing conceptions of the functions of the clinical psychologist have come about through his efforts to deal with these problems of the "Gaussian tail-enders" in more dynamic ways. There has been a shift in the climate of opinion from the desire to be scientific after the manner of the physical sciences, to an awareness of the importance of understanding human behavior and describing human relationships in terms of motive. Since the predicting of human behavior requires more variables than a regression equation can accommodate, must the clinical pyschologist become more of an artist than scientist?

THE CLINICIAN'S CHANGING ROLE

It has been the clinician's own increasing dissatisfaction with his role that, under pressure of service demand, has determined the direction of his growth and his changing status and professional stature.

It is understandable, in view of his antecedents, that the clinician should seek to develop new tests to assess more and more aspects of human abilities, that he should become especially interested in the identification and measurement of personality traits and in the little-understood problem of their interaction. How shall the dynamics of interpersonal relations and the ways of the unconscious determinants of behavior be taken into account in the clinician's task of dealing with people in trouble? If the clinician "now must do far more than produce a numerical IQ for the case record" [14, p. 4], or identify a schizophrenic by means of Rorschach signs, he must have more tested knowledge than is now available concerning the psychology of personality.

THE RISE OF PROJECTIVE TESTING

That the paper-and-pencil tests of personality have contributed so little to an understanding of the problems of human adjustment with which the clinician deals has been one of the reasons for the very rapid increase in the use of projective tests. Before the period of the second World War psychometric tests had mushroomed into existence by the thousands, but among all of these thousands only a relatively small number were being generally used by clinicians. In 1935 a list of tests in use in psychological clinics in the United States prepared by a committee of the Clinical

Section of the APA [22] includes some two hundred of the approximately four thousand listed by Buros for that period. Only twenty personality tests were mentioned in the entire list of tests in use in the clinics reporting, and none was included among the top ten arranged in order of frequency of mentions. A decade later comparable data were collected by Louttit and Browne [13]. In their 1946 survey the number of personality tests mentioned by at least 10 per cent of the clinics reporting had increased by two, and the Rorschach and Thematic Apperception Test are included among the first five on the list.

Other reasons for the rapid spread in the use of projective tests have had to do with significant changes that have been occurring in the clinician's self-concept and his changing perception of his role as his social responsibilities grow and expand. Projective tests have become important tools for the psychotherapists. These tests command the attention and respect of our collegues in the medical fraternity, the psychiatrists. They constitute moreover, the basic technological structure upon which is being built a new systematic point of view, projective psychology with its own theory of personality. This new projective psychology has been aptly characterized as a psychology of protest. As both behaviorism and Gestalt psychology came about as protests against the established psychologies called structural, so this emerging projective psychology runs sharply counter to the traditions that have characterized individual psychology in America. Having something to push against, it can move.

Objectivism has long been the hallmark of scientific respectability. The deliberate rejection by projectivists of demands for "observer reliability and public availability of operations" seems, indeed, like "an act of scientific irreverence" [2, p. 127]. It has been pretty generally agreed that the traditional measurement methods have been ill adapted to meet the requirements of projective theory upon which the new techniques have been based. Fortunately, however, the lack of methods suitable for testing dynamic personality theory has not led the majority of clinicians to discard all attempts at validation. There are many indications that the cultist view of Rorschach untouchability is yielding to a determined effort to devise new methods for justifying the clinician's faith in one of his most useful devices for understanding the whole person. Witness the current adaptations of the techniques of the analysis of variance by Block, Levine, and McNemar [3], and their application by Cass to the quantitative study of Rorschach patterns as related to clinical diagnosis [6].

THE DEMAND FOR THERAPY

New service demands, especially in periods of emergency, have always, in clinical psychology, had a marked effect in determining the direction of growth. In 1917 the attempts to assess the *abilities* of men established a new era of intelligence testing. In 1945, after four years of war, we found ourselves facing a new demand, the demand for therapy. Here again and more sharply than ever before, we have been faced by basic conflicts between our ancient tradition as scientists and the inescapable urge to render effective social service. Does one assess hunger or feed it? What does a humane scientist do about that? Or is a humane scientist a contradiction of terms? The outstanding development of clinical psychology's last decade has been the enthusiastic reception by the clinician of responsibility for therapy. Now he needs more basic facts and surer methods, and, though necessity be the mother of invention, gestation takes time.

There are now new and promising techniques for therapy. The year 1942 marked the publication of Rogers' *Counseling and Psychotherapy* and Frederick Allen's *Psychotherapy with Children.* There are the new projective tehniques and the rationale of Freudian and other dynamic theories of personality. We need time to acquire new skills for clinical practice, but even more importantly we need criteria for correct diagnosis, more tested knowledge about what constitutes effective therapy, and less guidance by what Miller characterizes as "the rules of the intuitive clinician's thumb" [14, p. 6].

PROFESSIONAL RESPONSIBILITY

At the Boston meeting of the APA one of

the guest speakers, a Doctor of Medicine, discussed the profession of psychology from a medical point of view [9]. Pointing out that psychology is a young profession, he proceeded to apprise us of the delights of being young. One of these assets of our youth, which he noted, is the freedom to create the demand for what we can do, *after* we have learned how to do it. Perhaps, as is so often the case in youth, we have not adequately guarded this freedom. Certainly clinical psychology, just like medicine under similar social demands, has had to reverse this happy sequence. Clinical psychology, we keep hearing, has been vastly oversold. We are, in a way, caught short. That this mischance has come about largely through the pressure of events over which we have had no control does not change the fact that we now find ourselves in the embarrassing position of having to learn how to do what we are doing *after* we have assumed the responsibility for doing it.

Since professional growth entails commensurate responsibilities, it is clear that as clinical psychologists we must accept the task of determining the value and limitations of the tools which we have created but which we share with other professional groups. At least we can undertake to report with discriminating caution the results of diagnostic testing involving the use of unvalidated techniques, and thus promote ultimate validation. It may well be that eventually the clinical psychologist with his background of training in test construction, experimental design, and statistical methodology will make his most significant contribution to "correct diagnosis and effective therapy" in research.

Our current practice of making such extensive use of mental hospitals for training facilities required by our VA training programs has its dangers as well as its values. The obvious advantages, for trainees, of opportunities to work with medical specialists, especially with psychiatrists, will, in the long run, be nullified if the effect of such association is to produce a pseudo-professional group of maverick psychiatrists and unanalyzed analysts. Avoiding the seduction of service demands in the atmosphere of a mental hospital requires a stern puritanism undoubtedly fostered, though not always adequately, by early indoctrination in the basic fundamentals of scientific psychology.

CONTROL BY INFORMATIVE FEED-BACK

We have sought to point out the chief trends in the development of clinical psychology. Now a trend implies movement in a discernible direction. The direction of change is discernible is historical perspective of clinical psychology. Can we now, assessing the past, extrapolate the immediate future? Cybernetics makes use of the concept of control by informative feed-back. Can knowledge give as much control to the clinical pyschologist as it gives to the anti-aircraft gun?

What has been happening in clinical psychology has been something like what occurs when a new generator is thrown into circuit with others already operating. The new unit has to be brought into frequency and phase relationship with the others already on the line. If the new generator is a little too slow when the operator throws it onto the line, then the other generators speed it up. It may speed up too much and have to be slowed down, and thus it oscillates around the true goal, back and forth like any moving object, "seeking" (as we say) equilibrium. In a governing system (servomechanism) there is likely to be this "hunting" before settling down on a goal. This is what Wiener, in his principles of cybernetics, calls control by feed-back and oscillation. "Motion," he points out, "is regulated by some measure of the amount by which it has not yet been accomplished" [19, p. 116]. Some kind of "control by informative feed-back" is evident in the dynamic interplay between events in this rapidly changing field of clinical psychology. Movement towards our goals is somehow regulated by the amount by which it has not yet been accomplished. In short, we may oscillate in ignorance under the pull of need, hunting for the balance of wisdom, until ultimately we settle down into the equilibrium of the ultimate goal.

In our early efforts to assess human abilities we developed intelligence tests. They gave us too much testing of too limited an aspect of the total individual, for the clinician needed to take account of more aspects of the personality. We were too slow, but negative feed-

back got us reversed in another direction, the upswing of personality testing. At first the personality tests were developed along the psychometric lines that had seemed to be suitable for intelligence testing, but the amorphous state of personality theory did not provide constructs that offered suitable material for measurement. The personality factors lacked stability, making prediction along psychometric lines inappropriate. Nor was there any agreement as to what or how many variables constitute an adequate statement about personality. Personality testing yielded too little about too much; there were too many variables of too little significance for the assessment of the individual.

Then came the backswing away from quantification, the movement towards projective testing. Projective tests are no longer tests at all in the psychometric sense, and their proponents have, until recently, offered determined resistance to the pressures to make the techniques conform to the usually accepted canons of scientific respectability in the matter of verification of assumptions and procedures. There is, however, already abundant current evidence that the movement is again changing. The steersman knows he has gone too far and is shifting back as he hunts. More and more the psychodiagnostician is demanding quantified statements of the relationship between projective tests and relevant independent measures of component personality variables. Perhaps we can begin to discern where some of our goals lie.

From one standpoint the condition of clinical psychology in its present state of flux seems to spell uncertainty and confusion. Perhaps for our future good we need to be uncertain, for these very uncertainties contribute not to confusion but to clarification. Belief in what we are doing is not inconsistent with an attitude of skepticism towards our explanations of how and why we are doing these things. I am in favor, not of what Conant has recently characterized as sneering skepticism, but of the skepticism that views our present concepts and conceptual schemes as a basis for further operations—a dynamic in contrast to a static approach to professional problems. After all a human being is a hunting organism, and it is fun to have been wrong if you can forever use it as the reason for getting right. Nor should we deplore the fact that the need for therapy runs always ahead of achievement in research, so long as ignorance is coupled with negative feed-back to produce wisdom. The scientist ought to be glad for ignorance, since it gives him his license to practice. The therapist ought to be pleased with the demand for his service, for it gives him the right to find out how to do better. This merry-go-round of basic principle and useful practice is the activity cage which I am now in position to offer to all clinical psychologists for their health and enjoyment, or at least to those of us who are truly alive and ready for action. Give us time, and oscillation, and hunting, and eventually we shall show history how men learned one more new way to look after the welfare of men.

Received May 12, 1951.
Early Publication.

REFERENCES

1. ABT, L. E. A theory of projective psychology. In L. E. Abt and L. Bellak (Eds.) *Projective psychology.* New York: Knopf, 1950.
2. ALLPORT, G. W. The use of personal documents in psychological science. *Soc. Sci. Res. Coun. Bull.*, 1942, No. 49.
3. BLOCK, J., LEVINE, L., AND MCNEMAR, Q. Testing for the existence of psychometric patterns. *J. abnorm. soc. Psychol.*, 1951, **46**, 356–359.
4. BORING, E. G. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century, 1942.
5. BORING, E. G. Great men and scientific progress. *Proc. Amer. phil. Soc.*, 1950, **94**, 339–351.
6. CASS, W. A. Quantitative Rorschach patterns as related to clinical diagnosis. Unpublished Doctor's thesis, Stanford Univ., 1950.
7. DARLEY, J. G., AND WOLFLE, D. Can we meet the formidable demand for psychological services? *Amer. Psychologist*, 1946, **1**, 179–180.
8. GLUECK, S., AND GLUECK, ELEANOR. *Unraveling juvenile delinquency.* New York: Commonwealth Fund, 1950.
9. GREGG, A. The profession of psychology as seen by a doctor of medicine. *Amer. Psychologist*, 1948, **3**, 397–401.
10. HEALY, W. *The individual delinquent.* Boston: Little, Brown, 1915.
11. HOLMES, A. J. *The conservation of the child.* Philadelphia: Lippincott, 1912.
12. LOUTTIT, C. M. The nature of clinical psychology. *Psychol. Bull.*, 1939, **36**, 361–389.

13. LOUTTIT, C. M., AND BROWNE, C. G. The use of psychometric instruments in psychological clinics. *J. consult. Psychol.*, 1947, **11**, 49–54.
14. MILLER, J. G. Contemporary trends in psychodiagnosis. In Molly R. Harrower (Ed), *Recent advances in diagnostic psychological testing*. Springfield, Ill.: Charles C Thomas, 1950.
15. MYERS, C. R. Prediction in clinical psychology. *Canad. J. Psychol.*, 1950, **4**, 97–108.
16. RAIMY, V. C. (Ed.) *Training in clinical psychology*. New York: Prentice-Hall, 1950.
17. SCHNEIDER, L. I. Rorschach validation: Some methodological aspects. *Psychol. Bull.*, 1950, **47**, 493–508.
18. TERMAN, L. M. The mental test as a psychological method. *Psychol. Rev.*, 1924, **31**, 93–117.
19. WIENER, N. *Cybernetics*. New York: Wiley, 1948.
20. WITMER, L. Retrospect and prospect: an editorial. *Psychol. Clin.*, 1908, **2**, 1–4.
21. WOODWORTH, R. S. The future of clinical psychology. *J. consult. Psychol.*, 1937, **1**, 4–5.
22. Report of Committee of Clinical Section of American Psychological Association. *Psychol. Clin.*, 1935, **23**, 23–27.

SYMPTOM PATTERNS IN A GROUP OF MENTAL HOSPITAL PATIENTS[1]

J. R. WITTENBORN

YALE UNIVERSITY

THIS is one of a series [5] of studies preliminary to the development of a quantified method for multiple psychiatric diagnoses. It is proposed to devise a rating scale procedure whereby a psychologist, psychiatrist, nurse, or other competent observer may (*a*) rate the currently discernible symptoms of a mental hospital patient, (*b*) score these ratings, and (*c*) prepare a profile which would indicate the degree to which the patient's symptom manifestations resemble each of the various symptom patterns which may be shown to exist among mental hospital patients. Some of the refinements and extensions of research in clinical psychology which could possibly result from an employment of such a device include: (*a*) the validation of many diagnostic implications of psychological tests, (*b*) a validation of many prognostic uses made of test scores, (*c*) a quantified study of background factors which are associated with the appearance of certain symptom patterns, (*d*) a quantified study of the conditions under which symptom patterns may change, and (*e*) a quantified evaluation of certain therapeutic procedures. Although this brief list does not exhaust the possible uses of a quantified multiple diagnostic procedure, it is sufficient to indicate advancements which could accrue from the use of such a diagnostic device or procedure. The term "diagnostic" is here used in a limited, descriptive sense, and the "diagnostic" method proposed is not concerned with etiological, prognostic, or dynamic considerations.

In order for a diagnostic procedure based on rating scales to serve in the expected manner, certain precautions in the development of the rating scales must be observed. For the purposes of the present report, it should be sufficient to state the necessary precautions and to describe briefly the ways in which the precautions were observed in the preparation of the rating scales.

1. The rating scales must sample adequately the important symptoms in mental hospital patients.—The patient group for which the scales were intended comprises all mental hospital patients suffering from or threatened by the so-called functional psychoses in which toxic and neurological factors are not believed to be present. In order to determine the important symptoms and their observational basis, a large number of psychiatrists were interviewed [5]. The relevant literature also was screened in an effort to develop a satisfactorily representative sample of symptoms.

2. The rating scales must be restricted to currently discernible behavior so that their repeated use can reveal changes in the patient.—It is apparent, therefore, that the scales cannot be concerned with many important factors in the patient's background or general circumstances. If they were not restricted to currently discernible behavior, their appropriate use would be limited by the adequacy of the patient's social history. In the present series

[1]A Special U.S. Public Health Research Fellowship and the support of the Institute of Human Relations, Yale University, made the present research possible. The computational assistance of the Department of Astronomy, Yale University, and of the U.S. Coast Guard Academy are also acknowledged gratefully.

The symptom rating scales were constructed with the collaboration of the Department of Psychiatry, Yale Medical School. The ratings were provided through the generous cooperation of the following members of the staff at the Northampton Veterans Administration Hospital: Dr. Isidor Scherer, Psychologist; Mr. Clinton Sherman and Mr. Alphonse Sootkoos, Nurses; and Dr. Felix Vecchione, Dr. Julius Kaplan, Dr. Russell Wharton, and Dr. Sol Winsten, Psychiatrists.

This report has been reviewed by the responsible officers of the Veterans Administration and is published with their approval. The opinions and conclusions are the writer's and do not necessarily reflect the opinion or policy of the Veterans Administration.

of studies, this restriction is not completely ignored, and the use of the scales is supplemented with biographical and situational data sheets and with somatic data sheets.

3. The scales must provide ratings which are relatively independent of the insights and sophistication of the rater.—The use of the scales cannot be based on dynamic interpretations or any inferences which require a skillful use of some particular theory or consideration of prior events. This restriction greatly limits the psychological scope of the scales, but without it their appropriate use would be dependent upon the sophistication of the rater. In order to avoid the danger of inconsistent use which would follow if the scales were formulated by highly sophisticated psychiatrists who made extensive use of some particular theory or point of view, the psychiatrists who assisted in the formulation of the scales and who participated in the pretesting were carefully chosen. The psychiatrists varied greatly in point of view, general sophistication, type of experience, and amount of experience. The pretesting was done chiefly by young, relatively inexperienced psychiatrists.

4. The rating scales must provide ratings which are relatively independent of any bias or theoretical persuasion of the rater.—It is, of course, extraordinarily important that the ratings be as free as possible from the bias of the rater. The bias could be due either to the manner in which a psychiatrist views a particular patient or the manner in which he views patients in general. The halo effect is ever present, and it would be presumptuous to say that it had been eliminated. Efforts were made to minimize it in the present study by employing rating scales for symptoms which were direct in their manifestations and required little judgment or interpretation on the part of the rater; the scales refer primarily to observable behavior or direct verbalizations. They were so constructed that the rater is obliged to read a series of descriptive statements before he can recognize the exact nature of the symptom involved, i.e., the first statement refers to a nonpathological level of behavior, and through successive steps the scale approaches a pathological extreme which usually refers rather clearly to some well-known and commonly observed symptom. None of the scales was labeled, and they are arranged in random order. Through these precautions it is hoped that the rater is confronted with a situation wherein he checks the descriptively appropriate behavior for his patient and is not constantly aware that he is rating some particular, labeled symptom.

5. The rating scales must be of such a form that each rating scale can be checked for every patient.—One of the factors which had to be considered in determining the form of the scales was the expectation that each scale would be correlated with every other one. In order to avoid the inconvenience of basing the different correlations on samples of different size (and also in order to make certain that each type of behavior was drawn to the rater's attention whether he had considered it important in his patient or not), it was necessary to devise scales which could be checked for every patient regardless of the applicability of the symptom for some patients. The problem of the applicability of the scale was handled by instructing the rater to check in each scale the item comprising the most extreme statement which he considered to be applicable to the behavior of his patients during the rating period. If the behavior of a patient was not noticeable in the respect referred to in the scale, the rater was instructed to check the first statement in the scale, that is, to check the item which is least extreme with respect to some particular pathological manifestation.

6. The rating scales must be feasible insofar as they draw upon behavior features which are likely to be known or may be readily ascertained.—A major factor in the feasibility of using a set of scales is content. Inasmuch as the scales were based on interviews with psychiatrists, the content may be reasonably regarded to be of general interest to psychiatrists and to refer to the kind of information which a psychiatrist is likely to have at his disposal.

7. The rating scales must be simple in form, and economical and convenient to use.—In the present study most of the scales comprise four items; none comprises more than five or less than three. It is also necessary to incorporate as few scales as possible in the total set. If a rater is asked to consider too many scales, he is likely to be indifferent, if not actually resistant, to the task. The set of scales employed in the present analysis can be rated in less than fifteen minutes by a psychiatrist who knows his patient well.

8. The rating scales must be in a form which does not bias the result of any analysis based upon them.—The form of the scales is critical in respects additional to those already discussed. For example, it is quite possible that (by forcing into one continuum manifestations which truly comprise two different continua) bipolar scales would obscure some of the clustering effects intrinsic to the manifestations of mental hospital patients. Bipolar scales are appropriate only when it is known that the antithetical forms of behavior at the extremes of the continuum bear a perfect negative correlation with each other. The mutually exclusive behavioral manifestations which this implies are rare in mental hospital patients; the use of bipolar scales would beg the question concerning the nature of symptom-patterning which the present investigation seeks to answer by relatively rigorous, empirical means.

9. The rating scales must be reliable.—It is not easy to check the reliability of rating scales which psychiatrists are to use in describing their patients. If the same psychiatrist uses the same scale on subsequent days, there is danger that a memory factor will produce a spuriously high reliability. If, how-

ever, he waits a substantial period of time before reapplying the rating scale to his patient, there is danger that the reliability may be spuriously lowered because the patient may have changed. The best type of reliability study would be one wherein two different psychiatrists rate the same patient independently for some specified period of observation. Since it is not often that two psychiatrists are studying the same patient at the same time, this type of reliability study is possible for only a very small number of patients; the limited data of this type which are available indicate that the scales could be highly reliable. Inasmuch as the scales were elaborately pretested and the limited reliability studies were favorable, it was believed that the reliability of the scales would be sufficient for the present purposes. The high correlation coefficients revealed when the scales were intercorrelated confirmed this supposition. Obviously, reliability studies will not indicate the sources of unreliability; for the present scales reliability may be limited not only by the form of the scales, but also by the reliability of the observations and by the interest of the rater.

SYMPTOM CLUSTERS OR SYNDROMES

The development of a quantified method for multiple psychiatric diagnoses implies the existence of clusters of symptoms. It is the purpose of the present report to describe a factor analysis of the matrix of intercorrelations resulting when each of the rating scales is correlated with every other one. There are several hazards which confront an investigator undertaking an exploratory study of this kind. It is quite possible that the behavioral symptoms of psychiatric patients do not cluster in a discernible manner, i.e., it is possible that in a group of patients the appearance of a particular symptom is no more likely to be associated with the appearance of another specified symptom than it is with the appearance of any randomly selected symptom. Nevertheless, inasmuch as psychiatrists have long spoken confidently of descriptive diagnoses and other investigators have found evidence of symptom clusters [1, 2, 3, 4], it seemed most likely that behavioral symptoms would cluster together in some consistent manner.

Presumably, the clustering of behavior symptoms revealed by a statistical analysis would resemble the clusterings implied by some of the diagnostic syndromes to which psychiatrists refer. There is no guarantee that such syndromes truly exist, however. It is quite possible that some of the symptoms which psychiatrists consider to form a syndrome are intrinsically unrelated to each other. Nevertheless, if psychiatrists believe consistently that certain symptoms go together, the ratings which they make for their patients may reveal this pattern of belief concerning symptom-clustering. Accordingly, if syndromes are revealed among the symptoms, the possibility remains that the syndromes are more descriptive of consistencies which exist in the behavior of psychiatrists than they are descriptive of consistencies which exist in the behavior of patients. As indicated in another paragraph, however, the rating scales are constructed and presented in a manner calculated to reduce the effect of any systematic bias among psychiatrists.

There are controls which may be employed to reveal evidence for possible differences in systematic bias between individual psychiatrists, and these controls will be described in a subsequent report. If it can be found that there are no differences in the symptom patterns revealed in the ratings by different individual psychiatrists, we need not concern ourselves too greatly with the possibility that a hypothetical pattern of bias characteristic of all psychiatrists determines the essential nature of any pattern found among symptoms. It must be remembered that the patterning among symptoms is not studied for its own sake, but merely as an essential step in the development of a dependable quantitative device for the description of patients' currently discernible behavior. If symptom patterns exist, either as a result of regularities in the behavior of the patient or regularities in the behavior of the psychiatrists, they may be used as economical points of reference for the description of the patient and some of the changes that occur in his behavior.

Inasmuch as it would be grossly uneconomical to attempt to describe patients in terms of a large number of symptoms, some sort of conceptualization of the patient's symptom manifestations is necessary for practicable description. Diagnostic syndromes are but constructs or inferences based upon the observation of clustering tendencies among a certain set of symptoms checked for a broad group

of patients. If a well-formulated, logically coherent theory for the psychoses were available, the intervening variables of this theory could provide the basis for economical description. Since no such theory is available, any phenomenal clustering tendency provides a plausible conceptualization and allows us to use some of the advantages of theory, i.e., measurement based on intervening variables. It is hardly necessary to remark that if such a theory existed there would be no motive for conducting the present study. It is possible, however, that a multiple quantified psychiatric diagnosis, although necessarily limited in its scope and naively descriptive in its nature, will serve as a useful tool in the kind of research that may be necessary for the development of the desired integrating theory.

The manifestation of symptom clusters, regardless of the exact conditions responsible for their appearance, takes us one step nearer to our goal, but there are other substantial hazards. For example, it is quite possible that the presence or absence of a clustering tendency among the symptoms, or the nature of any clusters revealed, may be determined by characteristics in the sample of symptoms which provides the basis for the rating scale. If the symptoms appropriate for a particular clustering tendency are omitted or insufficiently represented, the scope of any quantified method of psychiatric diagnosis based on the symptom-clustering is limited. Any fortuitous overrepresentation or underrepresentation of symptoms relevant to any particular cluster or clusters may result in a distortion of the sypmtom pattern revealed and could, as a consequence, delimit the usefulness of the quantified diagnostic procedure. Several devices were employed in order to safeguard the sampling of symptoms; noteworthy among these was continuous collaboration with a large number of psychiatrists, psychologists, and social workers, all with different backgrounds and interests; the published literature also received careful scrutiny.

It is equally evident that the presence or absence of particular behavioral manifestations in a patient group will affect the manner in which the behavioral manifestations are observed to cluster together. In a patient group there are many characteristics which conceivably could influence the manner in which symptoms cluster together. Among the more obvious possibilities would be age, sex, and education. The factors which are responsible for illness are likely to be important, and it is quite possible that the symptom patterns manifested by patients with atrophy or deterioration of the brain would be different from the symptom patterns manifested by patients whose illness appears to be unrelated to any neurological or toxic factors.

THE DATA AND THE METHODS OF ANALYSIS

The data on which the present analyses are based were provided by several psychiatrists and two psychiatric nurses. The subjects were 140 patients at the Northampton Veterans Hospital. Although the subjects were all males, they varied greatly with respect to age, educational attainments, social status, cause of illness, and presenting symptoms. Because of the heterogeneity of the patients and of the raters, the data are not ideal for the purposes of revealing consistencies in symptom patterns. If, however, there is present among mental hospital patients a consistent patterning of symptoms which is to an important degree independent of both the rater and the exact nature of the patient group, such symptom patterns are likely to be revealed in the present analysis. If clear-cut symptom patterns are revealed, one may assume that symptom patterning is an important phenomenon and be encouraged to study its precise nature minutely in homogeneous samples. If conspicuous clustering of symptoms is not revealed in a heterogeneous sample such as the present one, it is likely that the clustering of symptoms is a capricious or relatively unimportant phenomenon and would not provide the economy desired for a systematic study of the behavior of psychiatric patients.

The intercorrelations for each of the 55 rating scales are given in Table 1. (The labels or identifying phrases for the rating scales are taken from the most pathological level of each scale and are provided for the purpose of the present report. The scales as used were not labeled in any manner.) Almost all of the rating scales have large correlation coefficients with several of the other scales. Five of the scales had few ratings above the nonpathologi-

TABLE 1

CORRELATION MATRIX

	1	2	3	4	5	6	7	8
1. Acute Insomnia								
2. Ideas change with spontaneous rapidity	263							
3. Unjustified sexual beliefs*	(Omitted)							
4. Cannot banish obsessive thoughts	306	268						
5. Delusional belief that he is evil	261	—152		329				
6. Gives in easily to others	058	—114		078	340			
7. In almost constant movement	278	396		263	045	—121		
8. Unaware of the feelings of others	035	242		084	—075	150	132	
9. Use made of physical disease symptoms	186	000		106	—027	036	043	—107
10. Refuses to eat	253	—067		124	242	103	035	245
11. Deliberately disrupts routines*	(Omitted)							
12. Temper tantrums	154	216		136	—107	—358	150	—032
13. Avoids people	—040	—106		074	209	244	—139	277
14. Shouts, sings, and talks loudly	068	446		075	—236	—315	306	257
15. Behavior disrupted by phobias	171	—150		396	426	148	133	—191
16. Incontinent because of own negligence	187	225		023	—061	324	146	434
17. Engrossed in plans	—126	—083		—046	—082	—110	088	—284
18. Feelings of impending doom	296	—097		290	482	182	026	—102
19. Cannot resist compulsive acts	207	287		186	—090	—060	311	191
20. Exaggeration of ability and well-being	—078	296		025	—226	—211	235	072
21. Unable to stick to or carry out any plan	244	049		257	228	272	156	259
22. Cannot believe that he can be helped	230	—072		313	584	399	—023	—034
23. Fears others misunderstand him	000	063		247	125	—084	150	—124
24. Patient's thinking clearly delusional	000	374		293	035	—015	065	459
25. No organic basis for complaints	193	—085		085	—034	123	013	—040
26. Feels systematically persecuted	029	128		014	—152	—228	—030	080
27. Believes others influence him	—144	182		156	—125	—194	—100	100
28. Desperately distressed by his anxiety	350	—117		432	550	116	104	—184
29. Organic pathology with emotional basis	181	—119		122	137	026	063	—203
30. Rigidly orderly*	(Omitted)							
31. Dramatically attention-demanding	039	203		—028	—185	—235	174	—093
32. All overt activity is at a minimum	086	—211		—059	298	408	—167	189
33. Grandiose notions	—144	289		101	—179	—131	086	028
34. Does not believe he has a problem	—209	267		—191	—462	—198	013	400
35. Compulsive acts continuous	244	262		279	023	063	226	094
36. Great variation occurs in rate of speech	214	539		249	—046	108	241	332
37. Initiates physical assaults	022	235		120	—233	—298	295	309
38. Delusions of homosexual attacks*	(Omitted)							
39. Mood changes very frequent and abrupt	193	148		153	—105	—126	251	050
40. Has made attempts at suicide	302	—116		203	361	153	026	—026
41. Failures of affective response	040	052		077	086	321	014	300
42. No concern over physical handicaps	—049	—043		—048	021	029	000	—170
43. Cannot make decisions	178	—012		127	344	531	028	226
44. Opinions exceptional to physical laws	—144	036		010	—143	—394	043	—057
45. Hallucinations	051	211		081	—033	110	132	393
46. Memory faults	153	268		033	059	275	200	374
47. Fear of committing an abhorred act	153	—101		150	333	022	106	—053
48. Words not relevant to recognizable ideas	078	290		105	051	293	135	363
49. Shows failure and blocking	244	—138		325	607	344	048	—060
50. Repudiates earlier insights	184	309		189	000	165	279	289
51. Speech is stilted	000	173		000	—120	—149	041	123
52. Overt homosexual demands*	(Omitted)							
53. Lies or steals	109	159		—089	—158	—084	277	175
54. Exaggerated affective expressions	120	210		068	—121	—173	318	096
55. Characteristically oppositional	106	093		000	—108	—113	221	366

*Omitted for statistical reasons; see text.

TABLE 1 (Continued)

	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1.																
2.																
3.																
4.																
5.																
6.																
7.																
8.																
9.																
10.	000															
11.	(Omitted)															
12.	076	—156														
13.	—124	305		—235												
14.	021	—017		335	—136											
15.	000	108		—105	137	—217										
16.	087	189		—101	206	095	—080									
17.	196	—160		137	—254	107	—054	—236								
18.	092	094		023	119	—176	363	—017	—023							
19.	—021	—070		138	097	082	103	322	—131	—026						
20.	000	—053		128	—271	353	—133	—024	197	—154	—108					
21.	—046	247		—185	332	—076	080	332	—285	181	091	—078				
22.	000	166		—111	356	—200	335	015	—124	437	—011	—239	222			
23.	201	000		211	—116	110	169	—118	342	147	—087	094	—014	012		
24.	—149	081		033	171	235	—023	261	—118	000	064	166	060	—026	086	
25.	575	000		048	—140	—066	083	126	073	215	013	—080	—101	063	084	—041
26.	054	—133		243	—105	296	—124	014	074	022	—131	348	—153	—220	210	318
27.	020	—016		158	115	227	014	—089	020	—059	—077	116	—120	—105	288	510
28.	124	118		082	121	—226	497	—078	—021	666	081	—243	250	426	240	—151
29.	476	069		098	—145	—046	—019	—063	229	134	—047	—024	—034	104	081	—205
30.	(Omitted)															
31.	240	—156		393	—419	419	—280	—054	489	—029	—079	362	—160	—177	432	041
32.	—021	382		—360	551	—314	088	272	—239	102	—167	—285	386	304	—136	155
33.	—025	—104		013	—149	223	—035	—154	286	—213	—085	471	—185	—148	163	285
34.	—165	087		—094	008	316	—365	124	024	—331	—020	277	—064	—430	—116	339
35.	034	000		—017	152	037	278	354	—137	040	655	—085	204	142	—039	101
36.	099	108		066	—021	399	—137	399	—068	—040	221	111	241	—035	135	313
37.	056	000		373	—090	440	—191	063	028	—215	216	140	—067	—191	145	167
38.	(Omitted)															
39.	120	073		402	—194	343	041	135	030	—036	133	075	054	—158	243	128
40.	—047	096		058	152	—141	177	—018	—120	336	078	—138	227	397	041	—040
41.	025	280		—207	429	—188	051	442	—150	—015	205	—185	371	039	028	158
42.	490	—125		137	—198	—050	—042	—023	250	—073	—085	128	—167	—065	036	—197
43.	017	263		—354	384	—232	186	344	—225	212	057	—298	533	233	—089	102
44.	000	—146		356	—132	335	—053	—192	234	—168	—142	321	—232	—094	326	230
45.	—157	180		008	183	137	—054	401	—144	—028	166	—039	247	—050	—009	432
46.	—147	172		—220	163	023	—087	284	—236	—075	117	—018	293	022	—110	262
47.	—048	—078		118	000	—130	163	—072	097	341	106	—120	115	226	194	—061
48.	—121	099		—168	203	013	—066	454	—123	—029	215	—020	161	013	—098	404
49.	065	088		—054	196	—237	283	000	—022	569	—048	—200	328	447	201	—028
50.	062	119		114	040	227	—071	219	021	—099	173	139	264	088	181	116
51.	074	—060		024	—016	228	—084	018	225	—029	—055	164	015	—156	043	200
52.	(Omitted)															
53.	045	000		157	—233	197	—186	237	—092	—188	175	226	109	—214	026	—039
54.	148	—101		234	—160	311	000	297	050	000	234	104	000	—235	144	244
55.	088	197		142	014	156	—181	280	—089	—209	146	165	199	—209	026	151

TABLE 1 (Continued)

	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1.																
2.																
3.																
4.																
5.																
6.																
7.																
8.																
9.																
10.																
11.																
12.																
13.																
14.																
15.																
16.																
17.																
18.																
19.																
20.																
21.																
22.																
23.																
24.																
25.																
26.	—022															
27.	—074	484														
28.	193	—136	—138													
29.	436	—132	—144	211												
30.	(Omitted)															
31.	045	461	074	—064	035											
32.	013	—110	—011	069	015		—358									
33.	—095	264	355	—288	—130		207	—198								
34.	—294	200	237	—500	—258		117	—046	289							
35.	042	—174	—069	145	—024		—147	—056	—067	—041						
36.	103	—062	085	000	—024		103	—037	154	081	203					
37.	—017	208	200	—164	—040		276	—245	074	085	000	215				
38.	(Omitted)															
39.	093	168	108	032	—022		298	—248	059	—082	105	335	367			
40.	—073	132	—048	315	—051		—029	181	—140	—266	000	—081	034		000	
41.	015	—161	013	105	—036		—229	556	—244	045	300	211	—177		—095	—015
42.	343	029	016	—016	513		134	—068	—041	—150	—054	026	022		024	—149
43.	000	—326	—192	191	—025		—317	534	—146	010	222	204	—293		—053	062
44.	—080	384	276	—178	—112		525	—311	237	117	—179	—022	295		278	—078
45.	—156	015	243	—051	—183		—049	191	083	159	179	352	102		244	—019
46.	—092	—188	—065	—113	—046		—105	314	028	122	184	471	000		049	—154
47.	—059	089	—074	409	000		178	053	—126	—277	—042	—082	034		000	464
48.	—090	011	062	—117	—158		—135	280	096	200	169	333	017		—019	—015
49.	094	—071	—033	628	142		—067	240	—201	—411	057	000	—172		—017	383
50.	—013	—049	021	022	106		078	034	069	050	164	430	165		242	025
51.	061	334	240	—092	—089		321	—014	342	165	—021	148	018		095	—194
52.	(Omitted)															
53.	—056	106	—069	—097	098		336	—174	—120	096	118	078	260		140	000
54.	031	186	115	053	018		308	—164	148	174	173	277	232		480	—181
55.	—027	113	068	—130	016		123	121	029	202	096	132	348		136	013

TABLE 1 (Continued)

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
1.															
2.															
3.															
4.															
5.															
6.															
7.															
8.															
9.															
10.															
11.															
12.															
13.															
14.															
15.															
16.															
17.															
18.															
19.															
20.															
21.															
22.															
23.															
24.															
25.															
26.															
27.															
28.															
29.															
30.															
31.															
32.															
33.															
34.															
35.															
36.															
37.															
38.															
39.															
40.															
41.															
42.	—020														
43.	454	—135													
44.	—342	041	—382												
45.	300	—125	214	—062											
46.	281	—100	435	—223	335										
47.	—030	—076	042	—063	—019	—078									
48.	308	—173	383	—208	315	488	000								
49.	124	000	334	—201	009	024	295	—054							
50.	199	116	156	000	289	398	—026	274	105						
51.	—016	—020	011	130	030	054	—061	123	—069	053					
52.	(Omitted)														
53.	115	215	—159	060	000	000	000	000	—127	219	—029				
54.	—016	039	044	312	222	—810	—122	109	—181	160	221		202		
55.	196	070	058	175	054	060	—054	083	000	202	140		436	282	

cal level or had rather small correlation coefficients with the other scales and for this reason were not included in the present factor analysis. It cannot confidently be stated whether the failure of these particular scales to show high correlation was due to a lack of reliability, a lack of intrinsic relationship with any of the other symptoms included in the present study, or due to characteristics of either the patients or the raters. It is interesting to note that all of the variables have a large number of small correlations, and several of them have important negative correlations. From this it is not possible to anticipate what the factor pattern may be, but it is obvious that there is no general factor, and it appears quite possible that some of the factors at least may be well defined and based upon rather distinct clustering tendencies.

Accordingly, the intercorrelations for 50 of the rating scales were submitted to a Thurstone centroid analysis. In the case of each centroid factor the reflections were made on the basis of inspection and the questionable sums were checked by trial additions; the factor loadings were computed by the use of an automatic calculating machine; in computing the products of the factor loadings, an IBM multiplier was used, and the sign changes and residual correlations were secured by an employment of the tabulator and summary punch. After seven factors were extracted, the residuals were quite small and revealed no clustering tendencies.

The centroid factor matrix was submitted to a lengthy series of orthogonal rotations, and the results of the orthogonal rotations are given in Table 2, the rotated factor matrix.[2]

RESULTS

From an arithmetical standpoint the results of this analysis are relatively clear-cut. Although it is not claimed that the clusters are uncorrelated, an orthogonal frame of reference describes them quite satisfactorily. In the orthogonal factor matrix shown in Table 2, there is relatively little overlap among the factors, i.e., for the most part symptom rating scales which cluster together to form one factor tend not to participate in the formation of another factor. There is some exception to this general independence, however. For example, several of the symptom rating scales which are important in the formation of Factor 1 also have some importance in the formation of another factor. Nevertheless, the clustering tendency among the symptoms included in the present analysis is remarkably clear and compares very favorably with the clustering tendency observed among highly reliable measures such as mental tests. The clearness of the clustering tendencies is not due solely to the patterning among the factors, however. Many of the factor loadings which are critical in the determination of a particular factor are relatively high and also compare favorably with the size of the factor loadings obtained in analyses of mental test data. Incidentally, the size of the communalities for a number of the rating scales is excellent evidence that the rating scales have a satisfactory degree of reliability; in general, it may be assumed that the reliability of a rating scale is at least as great as its communality.

The fact that the clustering tendency is relatively clear-cut and that the factor loadings are high is encouraging to the purposes of this investigation because it suggests that from an arithmetical standpoint it is possible to score an individual's set of symptom rating scales with respect to a limited number of factors.

The encouraging implications of this analysis are not limited to arithmetical considerations alone. The nature of the symptom clusters revealed may be related to certain prevalent notions concerning the manner in which psychiatric symptoms do cluster together. Although this fact may be interpreted as an evidence for the validity of certain types of psychiatric thinking with respect to descriptive diagnosis, this fact has great relevance for the purposes of the present study. It suggests that an analysis such as this one, although quite foreign to clinical procedures and beliefs, may nevertheless produce results which are clinically plausi-

[2]To reduce printing costs, the centroid factor matrix and the orthogonal transformation matrix have been deposited with the American Documentation Institute. Order Document 3240 from American Documentation Institute, 1719 N Street, N.W., Washington 6, D.C., remitting $1.00 for microfilm (images 1 inch high on standard 35 mm. motion picture film) or $1.00 for photocopies (6 × 8 inches) readable without optical aid.

TABLE 2
ROTATED FACTOR MATRIX

		1	2	3	4	5	6	7	H^2
1.	Acute Insomnia	101	19	179	*358*	*413*	− 95	−129	367
2.	Ideas change with spontaneous rapidity	62	196	− 40	*683*	−120	115	83	545
3.	Unjustified sexual beliefs	(Omitted)							
4.	Cannot banish obsessive thoughts	− 32	68	90	*382*	*480*	249	− 66	456
5.	Delusional belief that he is evil	− 50	−252	− 7	− 42	*731*	− 79	96	618
6.	Gives in easily to others	19	*−586*	35	72	238	−240	107	476
7.	In almost constant movement	198	250	91	*485*	143	− 61	− 62	373
8.	Unaware of the feelings of others	*518*	−248	−215	*305*	−122	264	81	560
9.	Use made of physical disease symptoms	3	− 1	*727*	28	22	− 47	54	535
10.	Refuses to eat	214	−291	− 50	74	165	96	− 16	175
11.	Deliberately disrupts routines	(Omitted)							
12.	Temper tantrums	173	*564*	174	68	111	130	− 51	415
13.	Avoids people	216	*−542*	−248	− 57	203	230	− 34	500
14.	Shouts, sings, and talks loudly	261	*455*	− 34	*334*	−171	191	196	492
15.	Behavior disrupted by phobias	−139	−100	48	− 34	*513*	120	−166	338
16.	Incontinent because of own negligence	*474*	*−410*	142	*336*	− 58	67	− 73	539
17.	Engrossed in plans	−226	266	*328*	− 11	− 29	− 12	*353*	355
18.	Feelings of impending doom	−211	−114	112	− 14	*693*	86	9	558
19.	Cannot resist compulsive acts	201	41	43	*417*	44	94	−538	518
20.	Exaggeration of ability and well-being	100	*364*	1	200	−219	3	*368*	366
21.	Unable to stick to or carry out any plan	*343*	*−364*	−126	264	*329*	−106	12	455
22.	Cannot believe that he can be helped	105	*−470*	281	102	170	94	*347*	479
23.	Fears others misunderstand him	− 27	268	247	99	259	171	249	302
24.	Patient's thinking clearly delusional	134	−124	−131	*334*	− 17	*620*	251	610
25.	No organic basis for complaints	−122	− 67	*631*	37	63	5	− 99	433
26.	Feels systematically persecuted	140	*302*	14	− 32	7	*356*	*437*	430
27.	Believes others influence him	− 12	87	− 4	46	− 78	*643*	299	519
28.	Desperately distressed by his anxiety	−141	− 38	187	16	*781*	25	−179	699
29.	Organic pathology with emotional basis	− 66	− 14	*610*	− 21	84	−216	−41	432
30.	Rigidly orderly	(Omitted)							
31.	Dramatically attention-demanding	123	*545*	252	100	− 3	− 48	*464*	603
32.	All overt activity is at a minimum	195	*−737*	− 26	− 78	180	− 61	174	654
33.	Grandiose notions	−216	226	− 82	256	−236	234	*419*	456
34.	Does not believe he has a problem	184	− 13	−215	144	−497	*325*	189	489
35.	Compulsive acts continuous	148	−124	80	*422*	108	100	−467	462
36.	Great variation occurs in rate of speech	180	− 43	85	*679*	− 42	80	117	524
37.	Initiates physical assaults	*414*	*441*	22	179	− 61	166	− 37	431
38.	Delusions of homosexual attacks	(Omitted)							
39.	Mood changes very frequent and abrupt	255	*341*	200	244	107	221	− 16	341
40.	Has made attempts at suicide	132	4	−156	− 54	*594*	−112	23	411
41.	Failures of affective response	279	*−611*	81	196	4	73	− 98	511
42.	No concern over physical handicaps	55	56	*621*	− 78	−130	−175	48	448
43.	Cannot make decisions	61	*−684*	− 18	244	219	− 84	39	588
44.	Opinions exceptional to physical laws	180	*558*	12	−135	− 38	264	289	517
45.	Hallucinations	236	−235	−112	*327*	− 20	*335*	23	344
46.	Memory faults	−110	*−350*	−321	*674*	−183	−220	− 44	776
47.	Fear of committing an abhorred act	− 35	103	− 57	4	*525*	−109	36	304
48.	Words not relevant to recognizable ideas	92	*−406*	−112	*456*	− 80	108	124	427
49.	Shows failure and blocking	− 50	−219	90	38	*685*	− 82	97	545
50.	Repudiates earlier insights	299	− 24	101	*502*	48	− 71	109	371
51.	Speech is stilted	− 35	58	83	187	−130	172	*359*	222
52.	Overt homosexual demands	(Omitted)							
53.	Lies or steals	*561*	219	147	105	−109	−223	− 51	459
54.	Exaggerated affective expressions	259	256	*311*	33	−131	*325*	9	353
55.	Characteristically oppositional	*625*	25	109	109	−100	63	63	433

ble. This confers certain advantages to the present research because it suggests that a quantified method for descriptive psychiatric diagnosis may not only have the advantage of being based on empirically determined syndromes, but it may also have the advantage of employing syndromes which are relatively familiar in the thinking of psychiatrists and other professional workers who are likely to use the rating scales.

The plausibility of these syndromes may be judged by the reader upon examination of Table 2. In the evaluation of the results practical significance is arbitrarily ascribed to all factor loadings as high as .300. Using this criterion, Factor 1 may be seen to be determined by a tendency for the following symptoms to occur jointly in individuals: unawareness of others' feelings, oppositional behavior, deceptive behavior, assaultive behavior, incontinence, and inability to stick to a plan. If a particular patient presented all these characteristics simultaneously, most clinicians would consider him to be schizophrenic. Some would say, "this reminds me of a hebephrenic schizophrenic." Others might say, "This reminds me of some deteriorated patients that I have seen, schizophrenic patients who have been in the hospital a long time and are past the excited, fulminating phase of their illness." Accordingly, one is tempted to consider Factor 1 as due to the presence of individuals in the sample who are either hebephrenic or "deteriorated" (the word "deteriorated" is here used to refer to the somewhat flattened patient who has been in the hospital several years; it does not imply anything irreversible or irrevocable about the trend or status of his illness).

Perhaps one of the most interesting factors revealed in the present analysis is Factor 2, which is quite reminiscent of Kraepelin's classical distinction between patients who are manic and patients who are depressed. There is a bipolar arrangement of the symptoms with respect to Factor 2. Certain ones cluster together at one extreme of what may be conceived as a manic-depressive continuum whereas other symptoms cluster together at the opposite end of this continuum. Since Factor 2 actually involves two different clusters, each of the clusters is discussed separately. The fact that the symptoms for one of the clusters have positive signs and those for the other clusters have negative signs has no particular significance, and the solution would have no different implications had the respective signs been reversed. One cluster comprises the following symptoms: temper tantrums, loudness, attention-demanding behavior, assaultive behavior, fantastic thinking, exaggerated sense of well-being, feelings of persecution, abrupt mood changes. This combination of symptoms is quite suggestive of those manifested by patients in manic states. As a consequence, this portion of Factor 2 is taken to suggest that a scoring continuum for manic behavior would be a reasonable and proper part of a quantified diagnostic procedure.

The other cluster in Factor 2 includes the following symptoms: gives in easily to others, avoids people, is incontinent, is very inactive, shows little affective response, can't make decisions, words not relevant to recognizable ideas, can't stick to a plan, and has memory faults. This combination of symptoms is quite characteristic of profoundly depressed patients.

Factor 3 comprises for the most part conversion-type symptoms, and the clustering tendency is made up as follows: makes use of physical disease symptoms, has no organic basis for his complaints, presents organic pathology which has an emotional origin, shows little concern over his physical handicaps, is engrossed in plans, and has exaggerated affective expressions. The combination of conversion-type symptoms with a "belle indifférence" and exaggerated affect is strongly suggestive of the type of patient diagnosed as hysterical or conversion hysteria. It is not surprising that this syndrome should be manifest in a sample of mental hospital veterans.

Factor 4 is determined by the following symptoms: spontaneous change in ideas, constant movement, compulsive acts, variation in rate of speech, memory faults, unrecognizable use of words, repudiation of earlier insights, insomnia, obsessions, unawareness of others, loudness, incontinence, delusional thinking, and hallucinations. The pattern of behavior implied by this group of symptoms suggests disturbed patients such as those suffering from catatonic excitement. This kind of excited or fulminating behavior is usually well represented among groups of mental hospital pa-

tients. The importance of such a factor is likely to vary with the nature of the sample, however, i.e., in a sample of patients who have been hospitalized for a long time, this factor may be less conspicuous than it would be in a sample of patients recently admitted to a mental hospital and in many cases in an initial fulminating state.

Factor 5 appears to be determined by individuals who are acutely anxious, or, stated more precisely, determined by individual differences in the degree of anxiety. This interpretation is based upon the occurrence of a cluster which comprises the following symptoms: insomnia, obsessive thinking, delusions of guilt, phobias, feelings of doom, subjectively experienced anxiety, suicidal attempts, fear of their own behavior, blocking, and inability to carry out plans. This cluster is suggestive of both acute anxiety states and patients in agitated depression. The depressive aspects of anxiety are familiar to all clinicians, and the distinction which the analysis makes between the acute anxiety state of Factor 5 and the profound depression described by Factor 2 is in itself of considerable interest.

Factor 6 describes a clustering of the following symptoms: delusional thinking, ideas of reference, persecutory ideas, denial of a personal problem, hallucinations, and affective exaggerations. The delusional features of this cluster are paranoid in their nature, whereas the hallucinations and exaggerated affect are suggestive of a schizophrenic process. The implications of this cluster may be likened to the common implications of paranoid schizophrenia.

Factor 7, like Factor 6, refers to a cluster of symptoms which involves persecutory ideas. It also involves attention-demanding behavior, grandiose notions, excessive planning, exaggeration of ability, and stilted speech. When this cluster is contrasted with the cluster that comprises Factor 6, it is seen that paranoia and exaggerated planning are present but that obvious delusions and hallucinations are absent. It would seem, therefore, that Factor 7 refers to a paranoid condition as contrasted with a paranoid schizophrenia for Factor 6.

DISCUSSION

The results of the present analysis have two kinds of implications. First, they provide a relatively scrupulous description of the manner in which symptoms cluster together in a sample of mental hospital patients. The clustering patterns revealed have direct implications for some of our thinking concerning descriptive aspects of patients' behavior.

The second and currently more interesting implication of this analysis is that a quantified method of description of patients may be devised in a manner to indicate the degree to which individuals in their symptom manifestations resemble stereotypes revealed by a relatively objective method. Before any general relevance may be claimed for a given set of syndromes, there must be evidence for the pertinence of these syndromes for more than one particular situation. Although this cannot be explored exhaustively, it is obviously desirable to show that a symptom patterning is not due primarily to local characteristics of the patients or of the raters in some particular mental hospital. It is necessary at least to repeat the investigation of symptom-clustering for another hospital wherein different patients and different raters are involved. The exact nature of the present symptom-clustering may be in part dependent upon the fact that the present sample includes males only, includes mostly patients who have been ill and have been hospitalized for more than a year, includes several old patients, and includes a number of patients whose illness does not appear principally and primarily psychological. For example, the sample includes a number of alcoholic patients and patients whose disturbance of behavior is believed to be caused by neurological defects due either to disease or injury. The old patients, the organic patients, and the patients who have been hospitalized for a long time, may be responsible for Factor 1. It is quite possible that such a factor would not be found in a sample of patients under 60 years of age suffering only from what is considered to be a functional disorder and for the most part in the initial fulminating stages of their illness. It is possible that in such a sample the deteriorated cluster would not be apparent and that the excited cluster would be more conspicuous than it is in the present sample. This possibility is important because if studies such as the present one were to provide

the basis for a quantified multiple diagnostic procedure, such a procedure would be most useful if it were adapted principally for evaluating the early states of the illness of mental hospital patients. This emphasis is suggested because it is the relatively young, recently ill, functional psychotic who most often is in critical need of the diagnostic and evaluative skills of the psychiatrist and the psychologist.

CONCLUSION

This is one of a series of studies preliminary to the development of a quantified method for a multiple psychiatric diagnosis. The present report describes the development of a set of symptom rating scales for mental hospital patients. These scales were rated for a heterogeneous sample of veteran patients. The scales were then intercorrelated and the intercorrelations were factor-analyzed. As a result of this analysis seven clusters or groupings of symptoms were revealed. The clusters are relatively clear-cut and clinically plausible. This is taken as evidence that it may be possible to score the ratings for a given patient with respect to each of the clusters, and on the basis of such a scoring prepare a profile which would indicate the degree to which the patient's symptom manifestations resemble each of the various symptom clusters existing among mental hospital patients.

Received November 13, 1950.

REFERENCES

1. KRAEPELIN, E. *General paresis.* (Monogr. Ser., No. 14.) New York: Nerv. ment. Dis. Pub. Co., 1913.
2. KRAEPELIN, E. *Dementia praecox.* Edinburgh: E. & S. Livingstone, 1919.
3. KRAEPELIN, E. *Manic-depressive insanity and paranoia.* Edinburgh: E. & S. Livingstone, 1921.
4. MOORE, T. V. *The essential psychoses and their fundamental syndromes.* Baltimore: Williams & Wilkins, 1933.
5. WITTENBORN, J. R. A new procedure for evaluating mental hospital patients. *J. consult. Psychol.*, 1950, 14, 500–501.
6. WITTENBORN, J. R., AND METTLER, F. A. Some psychological changes following psychosurgery. *J. abnorm. soc. Psychol.*, in press.

THE EINSTELLUNG TEST OF RIGIDITY: ITS RELATION TO CONCRETENESS OF THINKING

ABRAHAM S. LUCHINS
McGILL UNIVERSITY

CONTEMPORARY literature indicates a widespread interest in the study, analysis, and measurement of rigidity of behavior and rigidity of personality. A considerable number of the current studies of rigidity [e.g., 1, 2, 5, 6, 7, 10, 11, 17, 19, 20, 21, 23, 24] employ procedures which might be described as *Einstellung* tests of rigidity, where *Einstellung* refers to the tendency to continue using an oft-repeated mode of solution in subsequent problems which are capable of solution by other, generally more simple methods [see 13, 14, 17].[1] (An Einstellung test of rigidity is described in the section on procedure of the present report.) Constructs and generalizations which stem from other tests or other criteria of rigidity—e.g., extent of variability of behavior [8], degree of satiation or cosatiation of an activity [12], inferences concerning rigidity drawn from responses to various clinical tests [22]—are being applied, at times rather uncritically, to "rigid behavior" as manifested in an Einstellung test. This is one of a series of papers reporting on investigations which are concerned with the applicability of such constructs and generalizations to behavior displayed in an Einstellung test of rigidity.

PROBLEM

On the basis of clinical observations and some experimental findings some psychologists have maintained that *concreteness of thinking* is a predeterminant of *rigidity*, and that the two tend to be positively related [e.g., 9, 26]. This relationship has been assumed to hold for rigidity as measured by an Einstellung test [23, 24]. Actual experimental demonstration of the relationship between concreteness of thinking and rigidity as measured by an Einstellung test has not yet been reported.[2] The problem of the present study is to investigate this relationship by using responses to a test of abstraction as the basis for evaluating the extent of concreteness of thinking.

SUBJECTS AND PROCEDURE

The subjects were 91 female freshmen students of Royal Victoria College of McGill University who had volunteered to serve as subjects. They had had no formal courses or training in psychology. To the group as a whole we administered the test of rigidity and, immediately afterwards, the test of abstraction to be described below.

Test of Rigidity. In order to study behavioral rigidity we employed one of the Einstellung tests which have been described in detail in our manual on rigidity [17]. The *S*s

[1] An *Einstellung* may be defined as a set which immediately predisposes an organism to one mode of motor or mental activity.

[2] While Rokeach's investigations [23, 24] have an oblique bearing on this relationship, they are concerned primarily with the relationship between rigidity and ethnocentrism [23], and with the effects of different "perception time" on both rigidity and concreteness of thinking [24]. Accepting the assumption that rigidity is positively related to concreteness of thinking, Rokeach hypothesizes (*a*) that if ethnocentrism should prove to be positively related to rigidity, then the more ethnocentric individual should be more concrete in his mode of thought, and (*b*) that decreased "perception time" should result in increases in both rigidity and concreteness of thinking. This latter expectation is based on the hypotheses that "abstract thinking can occur only when there is time available to think through a problem" and that if abstract thinking does not occur then the "inevitable consequence . . . is behavioral rigidity" [24, p. 215]. Rokeach measures concreteness of thinking in terms of the mean number of words *S* employs in writing his responses to the problems of the Einstellung test and in terms of the frequency of use of scratch paper in solving these problems.

are given the task of figuring out how they can obtain a stipulated volume of water in each of a series of numerical problems. The problems are as follows:

1. Given: an empty 29-quart jar, an empty 3-quart jar; measure 20 quarts of water.
2. Given: an empty 21-quart jar, an empty 127-quart jar, and an empty 3-quart jar; measure 100 quarts of water.
3. Given: an empty 14-quart jar, an empty 163-quart jar, and an empty 25-quart jar; measure 99 quarts of water.
4. Given: an empty 18-quart jar, an empty 43-quart jar, and an empty 10-quart jar; measure 5 quarts of water.
5. Given: an empty 9-quart jar, an empty 42-quart jar, and an empty 6-quart jar; measure 21 quarts of water.
6. Given: an empty 20-quart jar, an empty 59-quart jar, and an empty 4-quart jar; measure 31 quarts of water.
7. Given: an empty 23-quart jar, an empty 49-quart jar, and an empty 3-quart jar; measure 20 quarts of water.
8. Given: an empty 15-quart jar, an empty 39-quart jar, and an empty 3-quart jar; measure 18 quarts of water.
9. Given: an empty 34-quart jar, an empty 85-quart jar, and an empty 17-quart jar; measure 17 quarts of water.
10. Given: an empty 26-quart jar, an empty 65-quart jar, and an empty 13-quart jar; measure 13 quarts of water.
11. Given: an empty 28-quart jar, an empty 76-quart jar, and an empty 3-quart jar; measure 25 quarts of water.
12. Given: an empty 14-quart jar, an empty 21-quart jar, and an empty 4-quart jar; measure 10 quarts of water.
13. Given: an empty 15-quart jar, an empty 32-quart jar, and an empty 6-quart jar; measure 9 quarts of water.
14. Given: an empty 35-quart jar, an empty 69-quart jar, and an empty 5-quart jar; measure 30 quarts of water.

The first task is for illustrative purposes. Problems 2 through 6 are the Einstellung or "set" problems, each of which can be solved by one relatively complex procedure. If the three jars in the order listed are designated as A, B, C, respectively, then the solution to each of these problems may be represented by the formula $B - A - 2C$; e.g., $127 - 21 - 2 \times 3 \doteq 100$ gives the solution to Problem 2.

These Einstellung or "set" problems are followed by four critical problems, each of which is solvable by the $B - A - 2C$ method as well as by one or two direct procedures. In Problem 7 the direct method may be represented by the formula $A - C$ $(23 - 3 = 20)$. In Problem 8 the direct method is $A + C$ $(15 + 3 = 18)$. In Problems 9 and 10 the direct methods are $A - C$ and the filling of the C jar, e.g., the ninth problem is solvable by any of these methods: $85 - 34 - 2 \times 17 = 17$ $(B - A - 2C)$, $34 - 17 = 17$ $(A - C)$, 17 (C).

The criticals are followed by four extinction tasks solvable by a direct procedure, $A - C$, but not by the complex procedure, $B - A - 2C$.

The problems are written on the blackboard one at a time. Each problem is allowed to remain two and one-half minutes and then erased. The *S*s are told to write their responses to each problem on a new sheet of the examination booklets which have been given to them, and not to turn back to any previous page. They are also told that they do not have to use all of the jars given in the statement of a problem in their solutions. [See 17 for precise wording of all instructions.]

After *S*s attempt a solution to the first, the illustrative problem, E indicates that it can be solved by filling the 29-quart jar and from it pouring off three times into the 3-quart container; this is also written on the blackboard in the form $29 - 3 \times 3 = 20$. After *S*s have been allowed two and one-half minutes for the second problem, E gives the solution verbally and by writing it in this manner: $127 - 21 - 3 - 3 = 100$. He also illustrates another method of solving this problem: $127 - 9 \times 3 = 100$. No further help is given.

We are interested in whether *S* employs the $B - A - 2C$ method or one of the direct methods $(A - C, A + C, C)$ in the criticals and whether he fails to solve the extinction tasks which are solvable by an $A - C$ procedure but not by the complex procedure. Solutions of the criticals in the $B - A - 2C$ method are usually regarded as indicative of the development of a mental set or Einstel-

lung, and failures of the extinction tasks are usually interpreted as evidence of difficulty in surmounting the set.[3] We have used failures of the extinction tasks as the criterion of behavioral rigidity [see 17]. Other investigators, e.g., Rokeach [23, 24], have used *B* — *A* — *2C* solutions of the criticals as the criterion of behavioral rigidity.

Test of Abstraction. As a test of abstraction we utilized the similarities subtest of the Wechsler-Bellevue Intelligence Scale. This choice was based on (1) its popularity as a standard clinical tool for studying conceptual behavior and (2) its suitability for administration in a group situation. According to Wechsler, responses to this subtest throw light "upon the logical character of the subject's thinking processes" [25, p. 86]. The items of the subtest consist of the following 12 pairs of words.

1. orange—banana
2. coat—dress
3. dog—lion
4. wagon—bicycle
5. daily paper—radio
6. air—water
7. eye—ear
8. egg—seed
9. wood—alcohol
10. poem—statue
11. praise—punishment
12. fly——tree

Each pair of words is presented, one at a time, in accordance with Wechsler's instructions [25, p. 177], with the exceptions that it is presented to a group rather than to an individual and that *S*s write their answers on paper instead of giving them orally. For each pair of words, *S*s are told to write in what way the members of the pair are alike.

Wechsler indicates that each response is to be scored either 0, 1, or 2, depending upon the degree and quality of the generalization, with the higher score going to "superior" responses and the lower ones to "inferior" responses [25, p. 86]. Specifically, 2 credits is to be assigned for the most fundamental likeness, 1 credit for any essential likeness, and 0 for failure to respond, mentioning a difference rather than a similarity, or mentioning some inconsequential similarity [25, p. 198]. Independently of one another, the writer and two other clinical psychologists scored all responses, using the criteria illustrations offered by Wechsler [25, p. 198] as the basis for evaluation. There was complete agreement among the three judges as to the score to be assigned to each response.

Rapaport [22] points out that there are roughly three different levels on which individuals seek the similarity between the pairs of words: the abstract conceptual, the concrete, the functional. On the abstract level of concept formation, a generic term which connotes the fundamental common characteristic generally links the two; e.g., wagon and bicycle (fourth test item) are both vehicles. On the functional level of concept formation, the two terms are generally linked by a function performed by the two "things" they represent or by a function which human beings perform with them; e.g., you ride on both the wagon and bicycle. On the concrete level, a specific common element of the "things" in question links the two; e.g., both the wagon and bicycle have wheels.

Examination of the criteria illustrations offered by Wechsler indicates that responses for which a score of 2 is allowed are generally generic terms which have been taken by Rapaport [22] and other clinicians as indicative of the abstract conceptual level, while responses for which a score of 1 is allowed are indicative of the functional or concrete level. In accordance with common clinical practice we therefore labeled all responses to which the judges had assigned a score of 2 as *abstract*, all responses to which they had assigned a score of 1 as *concrete-functional*, and all to which they had assigned a 0 score as *failure*. As a further check, we asked the two other judges to rescore the responses, this time using the italicized scoring categories. There was complete agreement among the two procedures, i.e., the rescoring did not result in any response being shifted from one category to another.

[3] In some cases *B* — *A* — *2C* solutions of the criticals and failures of the extinction tasks may not be the resultants of an Einstellung, but may be brought about by other phenomena, e.g., by various assumptions or interpretations on *S*'s part [cf. 14, 17].

TREATMENT OF DATA

After scoring *Ss*' responses on the Einstellung test of behavioral rigidity, we eliminated from further consideration the papers of 20 *Ss* who had failed to solve at least one of the set problems, Problems 2 through 6, or who had solved any of them in a manner other than the *B* — *A* — *2C* manner. This was done in order to keep constant the amount of practice in the use of the *B* — *A* — *2C* procedure prior to the presentation of the criticals. There then remained 71 *Ss* whom we attempted to group into four classifications in order to simplify the comparison of responses on the two tests. This resulted in the elimination of 8 *Ss*, so that there remained 63 *Ss* who fell into one of four groupings.

Group I consists of 25 *Ss* who solved all four critical problems and all four extinction tasks by direct methods, i.e., who did not show any *B* — *A* — *2C* solutions or failures of these problems.

Group II consists of 9 *Ss* who used the *B* — *A* — *2C* method in the first or second critical problem, or both, but who solved all subsequent problems by direct methods.

Group III consists of 15 *Ss* who used the *B* — *A* — *2C* procedure in the first three critical problems or in all four critical problems but who solved all four extinction tasks.

Group IV consists of 22 *Ss* who used the *B* — *A* — *2C* method in all four critical problems and who failed to solve any of the four extinction tasks.

Interpreting these groupings in the terminology of Einstellung, we may say that *Ss* in Group I showed no overt signs of having developed an Einstellung or "set" for the use of the *B* — *A* — *2C* method, that *Ss* in Group II showed signs of having developed an Einstellung from which they readily recovered, that *Ss* in Group III did not recover quite so readily from their Einstellung, and that *Ss* in Group IV apparently manifested no recovery from their Einstellung, which was so strong that it blinded them to the solution of the four extinction tasks.

If failures to solve the *extinction* tasks are taken as evidence of behavioral rigidity, then it may be said that the least and the most rigidity were manifested by the *Ss* in Groups I and IV, respectively.[4] If *B* — *A* — *2C* solutions of the *critical problems* are accepted as evidence of behavioral rigidity [see 23, 24], then it may be said that the amount of rigidity consistently increases from Group I to Group IV, since the percentage of such solutions in the critical problems increases from group to group.

RESULTS

Classification of *Ss* into four groupings was based on their responses to the test of rigidity. In order to compare these four groups for responses to the test of abstraction, we determined the percentages of *Ss* in each group whose replies to a given test item of the similarities subtest fell into a particular scoring category, i.e., concrete-functional, abstract, or failure. The results are contained in Tables 1, 2, and 3.[5]

[4]Solution of the critical problem in the *B* — *A* — *2C* manner may or may not be followed by failures of the extinction tasks; failures of the extinction problems are almost invariably preceded by *B* — *A* — *2C* solutions (or occasionally, failures) of the critical problems. Thus, the criterion of behavioral rigidity based on the extinction problems actually entails the criterion based on the critical problems, while the reverse is not necessarily the case. Hence the most rigidity, based on the former criterion, would involve failures of all extinction problems as well as no direct solutions of the critical problems (Group IV), while the least rigidity would involve no failures of the extinction tasks as well as direct solution of every critical problem (Group I). Because of this we decided to compare Group IV with Group I, rather than with the combined data of Groups I, II, and III, none of whose *Ss* failed the extinction tasks.

[5]It will be noted that we are presenting the results for each item of the similarities test, instead of dealing with scores on the test as a whole or with the percentage of the total responses which fell into a given scoring category. While the latter procedures would undoubtedly have simplified the presentation of data, it seems to us that they might have beclouded the meaning of the results. To begin with, a total score on the similarities test does not necessarily reveal the extent to which it is compounded of concrete-functional, abstract, or failing responses. For example, a score of 12 might be based on concrete-functional responses (scored as 1) to each item, or on abstract responses (scored as 2) to the first six and failures on the remainder, or on a variety of other combinations. Moreover, evaluation of responses to the similarities test on the basis of total score or on the basis of the percentage of responses which fell into a given category presupposes that the test items are rather homogeneous in nature so that, for example, an abstract re-

When responses to the extinction tasks are used as a criterion of behavioral rigidity, then the least and the most rigidity were manifested by *S*s in Groups I and IV, respectively. If rigidity, based on this criterion, is positively related to concreteness of thinking, based on responses to the similarities test, then the following results may be expected:

1. Group I should show a smaller percentage of concrete-functional responses than Group IV.

2. Group I should show a larger percentage of abstract responses than Group IV.

3. Group I should show a smaller percentage of failure or the same percentage of failure as Group IV, since a failure does not contribute positively to the total score and is sometimes interpreted as an "inferior score" and as evidence of inability to abstract the similarity between the two words.

On the whole, the results reveal a fairly consistent but unreliable trend in line with expectations.

TABLE 1
PERCENTAGES OF CONCRETE-FUNCTIONAL RESPONSES TO THE SIMILARITIES TEST

Group	*S*s	Test Items											
		1	2	3	4	5	6	7	8	9	10	11	12
I	25	8	16	4	24	64	44	24	16	24	24	40	16
II	9	0	11	11	22	67	33	33	11	11	33	33	0
III	15	7	33	7	13	60	33	27	27	13	20	33	0
IV	22	9	27	5	18	55	41	50	23	18	50	50	23

From the data contained in Table 1 it can be ascertained that for eight of the twelve items of the similarities test the results with regard to concrete-functional responses are in accord with expectation, i.e., Group I contributes a smaller percentage of such solutions than Group IV. The results which go against expectation are found in test items 4, 5, 6, and 9. If the 5 per cent level of confidence is used as the criterion of statistical significance, then none of the differences between Groups I and IV, with regard to concrete-functional responses, has a t-value which denotes significance.

TABLE 2
PERCENTAGES OF ABSTRACT RESPONSES TO THE SIMILARITIES TEST

Group	*S*s	Test Items											
		1	2	3	4	5	6	7	8	9	10	11	12
I	25	88	84	96	76	36	48	44	80	68	60	28	48
II	9	100	89	89	78	33	56	45	78	67	56	0	44
III	15	93	67	93	87	33	67	20	73	73	53	20	47
IV	22	91	73	95	82	32	50	14	73	45	32	14	32

As can be determined from Table 2, in nine of the twelve test items the results with regard to abstract responses are in accord with expectation; i.e., Group I contributes a larger percentage of such responses than Group IV. For test item 7 the difference is significant at the 5 per cent level ($t = 2.34$). Results which go against expectation are found in test items 1, 4, and 6, but none of the differences is significant.

TABLE 3
PERCENTAGES OF FAILURES ON THE SIMILARITIES TEST

Group	*S*s	Test Items											
		1	2	3	4	5	6	7	8	9	10	11	12
I	25	4	0	0	0	0	8	34	4	8	16	32	36
II	9	0	0	0	0	0	11	22	11	22	11	67	56
III	15	0	0	0	0	7	0	53	0	13	27	47	53
IV	22	0	0	0	0	14	9	36	5	36	18	36	46

From Table 3 it can be determined that in eleven of the twelve test items the results with regard to failures are in line with expectation. In items 2, 3, and 4, both Group I and IV have zero failure; in all successive items, Group I has a smaller percentage of failure. For test item 9 the difference is significant at the 5 per cent level ($t = 2.37$). It is interesting to note that only on the first test item, which some clinicians regard as one of the simplest on the list, does Group I show a

sponse to the first item may be regarded as equivalent to an abstract response to the last. Actually, it is not unusual in clinical practice to consider the items as not equal in difficulty and hence as not equally indicative of abstract thinking on *S*'s part. Rapaport [22], for example, divides the items into three groups consisting of an easy group (items 1, 2, 3, 4, 5, 8), an intermediate group (items, 6, 9, 10), and a difficult one (items 7, 11, 12), and advocates an item analysis. For these reasons it seemed to us that presentation of the results item by item would yield a clearer picture than would analysis in terms of total or average scores on the test as a whole.

larger percentage of failure than Group IV; the difference is not significant.

Thus, for each of the three categories the trend of results is in accord with expectation based on the assumption that behavioral rigidity (based on responses to the extinction problems of the Einstellung test) is positively related to concreteness of thinking (based on responses to the similarities test). While only two of the differences which are in line with expectation are significant at the 5 per cent level, none of the differences which go against expectation has a significant *t*-value.

Until now we have been discussing the results in terms of failure of the extinction problems. Solution of the criticals in the *B* — *A* — 2 *C* manner has also been used as a criterion of rigidity [23, 24]. If behavioral rigidity, based on this latter criterion, is positively related to concreteness of thinking, then the following should be expected:

1. The percentages of concrete-functional responses should increase from Group I to Group II to Group III to Group IV.

2. The percentages of abstract responses should decrease from group to group.

3. The percentages of failure should increase or at least not decrease from group to group.

Examination of Tables 1, 2, and 3 reveals that in each scoring category the trend of results does *not* support this expectation. In not one of the twelve test items does the percentage of concrete-functional responses (Table 1) consistently increase from Group I to Group IV. In only one test item (item 10) does the percentage of abstract replies (Table 2) vary according to expectations. In only four test items does the percentage of failures (Table 3) remain constant (items 2, 3, 4) or increase (item 5) from group to group. Not one of the differences between any two of the four groups—aside from the two significant differences already referred to in previous paragraphs—is significant at the 5 per cent level of confidence.

DISCUSSION

The difference in the trends of results when one rather than another criterion of behavioral rigidity was employed, raises again the problem of just what constitutes "rigid behavior" on the Einstellung tests [cf. 16, p. 455]. If rigidity is regarded as a lack of shift in behavior when the situation calls for a shift [cf. 23], then it may be said that a change from the "set" procedure (*B* — *A* — 2*C*) to a direct solution is called for in the extinction tasks but is not demanded in the critical problems, since the latter are solvable by the "set" procedure. It is on this ground that we have criticized the utilization of *B* — *A* — 2*C* solutions of the criticals as evidence of rigidity [16, p. 455].

In the present investigation we accepted the assumption that the underlying thought process can be inferred from a response to the similarities test item [cf. 22], e.g., we assumed that a response which was given a score of 2 and labeled as "abstract" was indicative of an abstract conceptual level. We also accepted the related assumption that increased concreteness of thinking would be manifested by increases in concrete-functional responses, decreases in abstract replies, and no decreases in failures. Actually, we recognize that overtly similar or even identical responses may be the resultants of thought processes which differ in their logical character, while overtly dissimilar responses may be the resultants of similar thought processes, so that one may be on very shaky grounds when one infers the logical character of conceptual behavior from the response, the end product of the thought process. In the hope of obtaining a more direct understanding of the logical character of conceptual behavior and of its relationship to behavioral rigidity, we are now conducting individual experiments in which the concern is not only with the final responses but with the actual thought processes which underlie (1) an individual's responses to the similarities subtest of the Wechsler-Bellevue Intelligence Scale (and to other tests of concept formation), and (2) the same individual's responses to an Einstellung test of rigidity.

It may not be out of place to refer here to what we consider to be the functions of our Einstellung tests. We believe that the examiner should not be primarily concerned with obtaining a "rigidity score" for *S* which may serve as a more or less invariant index of his

behavioral rigidity or which may be used to compare his performance with that of other *S*s under standard test conditions. Rather, we believe that test conditions themselves should be treated as a variable to be manipulated experimentally. The examiner should attempt systematically to vary test conditions for a particular *S* in order to obtain changes in his behavioral rigidity, e.g., in order to develop an Einstellung in an *S* who previously showed none or vice versa, in order to weaken or strengthen an Einstellung or recovery from it. Such systematic variations may yield an understanding of the functional relationship between test conditions (or more broadly, field conditions) and behavioral rigidity.

A similar approach may prove fruitful when applied to tests of concept formation. This would require the clinician to attempt to vary systematically the conditions under which he administers such a test to an *S* and to note the relationship between changes in test conditions and changes in *S*'s responses and conceptual behavior. It would mean that the clinician would not be exclusively concerned with gauging the logical character of conceptual behavior evinced in a *standard* test situation. It would also mean that the clinician would not be primarily concerned with classifying the conceptual level *characteristic* of *S* or *preferred* by *S*, or with classifying *S* himself, e.g., as a "concrete-minded" or an "abstract-minded" individual. Experiments conducted by the writer and his students, some completed and others in progress, indicate that such classifications are sometimes invalid since the conceptual level on which *S* appears to operate during the administration of a test may not be the same one on which he operates in other situations, e.g., an *S*'s responses to a similarities test and the apparent underlying conceptual behavior have been found to change with changes in initial instructions, changes in social atmosphere, etc.

Systematic variations of the conditions under which tests of rigidity and tests of concept formation are administered, may help to broaden our understanding of the relationship between behavioral rigidity and conceptual behavior in so far as they make possible the comparison of (*a*) the kind of conditions under which *S* reveals a certain extent of behavioral rigidity and the kind of conditions under which he reveals a certain extent of, say, concreteness of thinking; (*b*) the apparent susceptibility to change of *S*'s behavioral rigidity, on the one hand, and of the logical character of his conceptual behavior, on the other; and (*c*) the kinds of changes in test conditions which are successful or not successful in producing changes in the two variables.

SUMMARY

We described an investigation of the relationship between behavioral rigidity, based on responses to an Einstellung test, and concreteness of thinking, based on responses to the similarities subtest of the Wechsler-Bellevue Intelligence Scale. The Einstellung test employed consists of a series of numerical problems in which *S*'s task is to obtain a stipulated volume of water by means of all or any of the given containers which have stated capacities. After the illustrative problem there follow several "set" problems, each of which is solvable by one mode of procedure; if the three jars in the statement of the problem are labeled *A, B, C,* respectively, then the "set" procedure can be represented by the formula $B - A - 2C$. There follow critical problems each of which is solvable both by the $B - A - 2C$ method and by one or two direct methods, e.g., $A - C$. The final problems in the series are extinction tasks each of which is solvable by a direct method but not by the "set" procedure. The 63 *S*s whose results were used were divided post-experimentally into four groups, depending on the number of $B - A - 2C$ solutions of the criticals and the number of failures of the extinction tasks.

Responses to the similarities subtest were scored as 0, 1, or 2, and labeled as failure, concrete-functional, or abstract. We assumed that greater concreteness of thinking would be evinced by increases in concrete-functional responses, decreases in abstract responses, and no decreases in failures.

We have used failures to solve the extinction tasks as an operational index of behavioral rigidity. Other investigators have used $B - A - 2C$ solutions of the criticals as a criterion of behavioral rigidity. When the former cri-

terion was employed, the results yielded a rather consistent but generally unreliable trend toward a positive relationship between behavioral rigidity and concreteness of thinking. When the latter criterion was employed, the results failed to yield any consistent trend.

In the discussion section we compared the two criteria of behavioral rigidity. We then questioned the validity of the assumptions which underlie our usage (and prevalent current usage) of the similarities test. Finally, we advocated systematic variations of the conditions under which tests of rigidity and tests of conceptual behavior are administered.

Received November 29, 1950.

References

1. Ainsworth, L. H. Rigidity and security. Unpublished Master's thesis, Univ. of Toronto, 1950.
2. Applezweig, D. Mimeographed outline of a revised thesis design. Wesleyan Univ., 1949.
3. Beier, E. G. The effect of induced anxiety on some aspects of intellectual functioning: A study of the relationship between anxiety and rigidity. Unpublished Doctor's dissertation, Teachers Coll., Columbia Univ., 1949.
4. Cattell, R. B. The riddle of perseveration: II. Solution in terms of personality structure. *J. Personality,* 1946, **14**, 239–267.
5. Christie, J. R. The effects of frustration upon rigidity in problem solution. Unpublished Doctor's dissertation. Univ. of Calif., 1949.
6. Cowen, E. L. A study of the influence of varying degrees of psychological stress on problem solving rigidity. Unpublished Doctor's dissertation. Syracuse Univ., 1950.
7. Cowen, E. L., and Thompson, A. W. Problem solving rigidity and personality structure. *J. abnorm. soc. Psychol.,* 1951, **46**, 165–176.
8. Fisher, S. Patterns of personality rigidity and some of their determinants. *Psychol. Monogr.,* 1950, **64**, No. 1.
9. Goldstein, K. Concerning rigidity. *Character & Pers.,* 1943, **11**, 209–226.
10. Harris, R. A. Personal communication concerning Doctor's dissertation, Harvard Univ., 1950.
11. Horowitz, L. Personal communication concerning Doctor's dissertation, New York Univ., 1950.
12. Kounin, J. S. Experimental studies of rigidity: I and II. *Character & Pers.,* 1941, **9**, 251–272, 273–282.
13. Luchins, A. S. The effects of Einstellung in learning by repetition. Unpublished Doctor's dissertation. New York Univ., 1939.
14. Luchins, A. S. Mechanization in problem solving. *Psychol. Monogr.,* 1942, **54**, No. 6.
15. Luchins, A. S. Proposed methods of studying degrees of rigidity in behavior. *J. Personality,* 1947, **15**, 242–246.
16. Luchins, A. S. Rigidity and ethnocentrism: A critique. *J. Personality,* 1949, **17**, 449–466.
17. Luchins, A. S. Examination for flexibility-rigidity of behavior. Mimeographed manual. Montrose, N. Y.: F. D. Roosevelt Veterans Administration Hospital, 1950.
18. Luchins, A. S. On recent usage of the Einstellung effect as a test of rigidity. *J. consult. Psychol.,* 1951, **15**, 89–94.
19. Meer, B. Personal communication concerning Doctor's thesis on rigidity and insecurity-security. V. A. Hospital, Perrypoint, Md., 1950.
20. Oliver, J. A. An analysis of some new tests of disposition rigidity. Unpublished Master's thesis, McGill Univ., 1950.
21. Pally, S. Personal communication concerning Doctor's dissertation. Univ. of Pa., 1950.
22. Rapaport, D., Gill, M., and Schafer, R. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1946. Vol. II.
23. Rokeach, M. Generalized mental rigidity as a factor in ethnocentrism. *J. abnorm. soc. Psychol.,* 1948, **43**, 259–278.
24. Rokeach, M. The effect of perception time upon rigidity and concreteness of thinking. *J. exp. Psychol.,* 1950, **40**, 206–216.
25. Wechsler, D. *Measurement of adult intelligence.* (3rd Ed.) Baltimore: Williams & Wilkins, 1944.
26. Werner, H. *Comparative psychology of mental development.* New York: Harper, 1940.
27. Werner, H. The concept of rigidity: A critical evaluation. *Psychol. Rev.,* 1946, **53**, 43–52.

THE PERFORMANCE OF YOUNG SCHIZOPHRENICS AND YOUNG NORMALS ON THE WISCONSIN CARD SORTING TEST[1]

ELIZABETH T. FEY
UNIVERSITY OF WISCONSIN

ALTHOUGH there are wide differences in the explanatory concepts utilized, there is general agreement that schizophrenic patients experience more difficulty than do normal individuals in the effective utilization of concepts or generalizations. This fundamental deficiency has been clearly established both through clinical observation and the application of structured testing techniques. A series of investigations of this problem [3, 6, 7, 11, 12, 22, 23, 24] was initiated by Goldstein's introduction of the sorting test technique and the provocative conclusions which he drew concerning the nature of the impairment of thinking in schizophrenia [7].

Although all of these studies have confirmed the presence of the general deficiency noted above, our understanding of the exact nature of this deficiency is still very limited. Goldstein's concept of an impairment of abstract attitude is provocative but poorly adapted to experimental validation, and the assumption of organic impairment which is implicit in the viewpoints of both Goldstein and Vigotsky finds far from universal acceptance today. Vigotsky's concept of regression from categorical to complex thinking is not only difficult to test, but is directly challenged by subsequent evidence that regression, in the sense of a peeling-off of successively acquired patterns of behavior, is not characteristic of schizophrenic illness [3, 4, 5]. Cameron's concept of faulty exclusion is a constructive step in the direction of reducing the observed deficiency to its component factors, but leaves many other steps to be taken.

At the present stage of research there is no general agreement concerning (1) the exact nature of the task presented by sorting test situations, (2) the characteristics of schizophrenic performance which militate against success, or (3) the factors inherent in schizophrenic illness which may be responsible for the difficulty. At least part of the confusion in this area can be traced to the utilization of testing techniques which are not adapted to objective scoring and quantitative analysis, and the absence of adequate normative data against which clear-cut comparisons can be made. The present study was designed to study the differences between the performance of schizophrenic patients and normal subjects on a sorting test problem. An attempt was made to overcome some of the difficulties mentioned through the utilization of carefully matched patient and control groups and a testing technique which is adapted to objective scoring and provides quantitative measures of many aspects of performance.

PROCEDURE

Description of the Wisconsin Card Sorting Test. The WCST is similar in its essential features to the Goldstein-Scheerer-Weigl type sorting tests [1, 8, 9, 10, 21].

The form of the test utilized in this study consists of a pack of 64 white response cards, each containing from one to four identical figures of a single color, and four stimulus cards. The four stimulus cards are placed on the upper half of a sorting tray divided into four compartments. From left to right, the stimulus cards contain one red triangle, two green stars, three yellow crosses, and four blue

[1]This article is based on a thesis submitted to the University of Wisconsin in partial fulfillment of the degree of Master of Science. The author gratefully acknowledges the helpful guidance of Dr. Ann Magaret, Dr. David A. Grant, and Dr. Norman Cameron.

circles. The 64 response cards represent all possible combinations of the four colors, four figures, and four numbers of figures on the cards, and can thus be sorted on the basis of color, form, or number. In the test situation, the board with a stimulus card placed above each of the four compartments in the order described was placed before the subject. He was given the pack of 64 response cards and the following instructions: "This is a matching test. You are to take the cards one at a time from the top of that pack and match each of them with one of these four cards [pointing to the stimulus cards]. Put each card in the box in front of the card you wish to match it with. After you place each card, I will tell you whether you are right or wrong. Now, go ahead, put the first card in one of the boxes, and I will tell you whether you are right or wrong." The instructions were repeated and wording altered when necessary to insure understanding, but no additional cues were given.

The experimenter followed a predetermined category order, and completion of the test demanded five shifts—from color, the initially correct category, to form, to number, back to color, and again to form and then number. In every category, the experimenter continued to judge as "wrong" any match which did not involve the correct category, until the subject had made ten consecutive correct matches. She then shifted the basis of her judgments to the next category, and the subject was given no other cue that any change had occurred. Every card placed by the subject throughout the test was recorded in terms of the properties matched, i.e., C, F, N, or any combination thereof. The test was concluded when the subject had successfully completed the six stages, or had sorted 64 cards in any one of the stages without making ten consecutive correct responses. At the conclusion of the test, each subject was asked to describe what he had been doing during the test, and his responses were recorded.

The WCST is easily and objectively scored in terms of the following measures:

1. *CR—correct responses.*
2. *PR—perseverative responses.* Perseverative responses are defined as any errors made by sorting on the basis of the previously correct category.
3. *NPE—nonperseverative errors.* Nonperseverative errors include all errors made upon some other basis than the previously correct category.
4. *UR—unique responses.* Unique responses are those which do not represent matches for any of the three major categories. For example, a response card with two yellow crosses, placed beneath the stimulus card with four blue circles.
5. *TE—total errors.* The sum of *PR, NPE,* and *UR.*
6. *TT—total trials.* This measure represents the total number of cards placed by the subject throughout the test, the sum of *CR* and *TE.*
7. *CC—categories completed.* This measure represents the number of stages of the test completed successfully, 0 to 6.
8. *CLR—conceptual-level responses.* Conceptual-level responses are defined as correct responses occurring in consecutive runs of three or more. The criterion of three or more consecutive correct responses was adopted as a measure of partial insight, since the probability of the chance occurrence of such runs is low, and examination of the records strongly suggests that they do not occur in a random way.

Description of the Groups. Twenty-two young schizophrenic patients and 47 young normal subjects were utilized in the study. The schizophrenic group contained 17 patients diagnosed "hebephrenic," 3 diagnosed "paranoid," and 2 diagnosed "simple." The groups were matched in age, tested intelligence, educational level, and ratio of men to women, as indicated in Tables 1, 2, 3, and 4. The schizo-

TABLE 1
RATIO OF MEN TO WOMEN

Group	Men	Women	Total
Patients	10	12	22
Controls	24	23	47

TABLE 2
AGE

Group	Range	Mean
Patients	16–33	28
Controls	18–34	26

TABLE 3
WECHSLER-BELLEVUE IQ

Group	Range	Mean
Patients	72–119	92.3
Controls	74–120	95.8

TABLE 4
EDUCATIONAL LEVEL (YEARS)

Group	Range	Mean
Patients	8–14	10.9
Controls	6–13	9.8

phrenic subjects were all patients in a state hospital,[2] and were selected on the basis of diagnosis, age, and testability. The control group was composed of production workers in two Madison industries and machinist-apprentices at an adult vocational school.[3] Every patient and normal subject who achieved a Full Scale IQ score of 69 or above on Form 1 of the Wechsler-Bellevue Adult Intelligence Scale was included in the study. Subjects scoring below 69 were excluded in order to avoid the special problems of mental deficiency.

It should be noted here that the requirement of obtaining an adequate Wechsler-Bellevue IQ score on all subjects utilized in this study produced a schizophrenic group which was, on the whole, very well organized and cooperative. Nineteen of the 22 patients studied had been hospitalized for less than ten months prior to testing, and 50 per cent of the total group were released to their families in the course of the following year. Only subjects whose responses were coherent and relevant and, hence, easily scored in terms of standard criteria were included in the group. In the previous studies cited, there was no attempt to obtain an IQ rating on the psychotic subjects, and they were, on the whole, older and more seriously disturbed than the present group.

Testing Procedure. Every subject included in the study was tested on Form 1 of the Wechsler-Bellevue Adult Intelligence Scale and the Wisconsin Card Sorting Test by the same examiner. The tests were administered in that order in a single testing session which averaged about two hours in length.

RESULTS AND DISCUSSION

Comparison of the Groups in Terms of Gross Measures of Success. On gross measures of degree of success achieved, the WCST differentiated clearly between the groups. Sixteen of the 22 schizophrenic subjects failed to complete the six stages of the test successfully, while only 8 of the 47 control subjects failed to do so. A chi-square test based on the number of subjects in each group passing and failing the test was significant at better than the .01 level of confidence. The difference in mean Categories Completed between the two groups, utilizing Student's *t*-test, was also significant at better than the .01 confidence level. The above results are particularly interesting in view of the fact previously noted, that the schizophrenic patients were young and exhibited relatively little disorganization. The WCST is apparently sensitive to factors present in the early stages of schizophrenic illness.

Relationship of Wechsler-Bellevue Scores to WCST Performance. Pearson product-moment correlation coefficients based on the score Categories Completed and Full Scale IQ scores revealed no significant correlation between tested intelligence and success on the WCST in either the control or patient group. The correlations obtained were .135 for the control group and .182 for the patient group. Nor was any significant correlation obtained between Categories Completed and any of the Wechsler subtests for either of the groups.[4] The latter result may have some interesting implications. It is often assumed that the inadequate performance of schizophrenic patients on problems of this nature may be attributed primarily to a disturbance in thinking. If so, we might expect such a disturbance to be reflected with some consistency in performance on some of the tasks utilized in standard intelligence tests. Specifically, in terms of Goldstein's analysis, we might expect special difficulty with block designs, object assembly, arithmetic problems, digit symbol, or similarities, since Goldstein considers the abstract attitude to be essential to solution of these types of problems. In so far as we can rely upon the statistical analysis, however, no such relationship appears to hold. The WCST apparently measures something different from

[2]The author wishes to express her appreciation to the staff of the Mendota State Hospital for their cooperation in pursuing this research.

[3]The author wishes to express her appreciation to the officers of the Celon Co. and the Oscar Mayer & Co. and the staff of the Adult Vocational School for their cooperation in pursuing this research.

[4]Although these results are suggestive and are reported for that reason, they should be interpreted with caution. The application of the Pearson product-moment correlation analysis to these data is not strictly valid in view of the restricted range of the score *CC* and the markedly skewed distribution of this score in the control group.

the factors assessed by standard intellectual tasks. In general, it seems clear that intelligence, as determined by standard testing techniques, is not a major factor in determining the adequacy of performance on the WCST within the IQ range tested.[5]

Comparison of Patients and Controls Successfully Completing the WCST. In the detailed analysis of the performance of the control and schizophrenic subjects on the WCST, three major comparisons were made: (1) between the 39 control and 6 schizophrenic subjects who succcessfully completed the six stages of the test, (2) between the total control and patient groups, including both "passes" and "failures," and (3) between the 8 control and 16 schizophrenic subjects who failed to complete the six stages of the test. The results for the three comparative groups of the application of Student's *t*-test to all of the scores utilized are presented in Tables 5, 6, 7, and 8.

TABLE 5

COMPARISON OF PATIENTS AND CONTROLS SUCCESSFULLY COMPLETING THE CARD SORTING TEST

Source	df	Means		t	Confidence Level*
		Patients	Controls		
TE	43	44.66	29.72	2.23	.05
PR	43	32.66	17.20	3.49	.01
NPE	43	18.16	13.58	1.29	—
UR	43	2.33	1.28	0.88	—
CR	43	91.16	79.89	1.94	—
TT	43	137.50	109.36	2.37	.05
CLR	43	17.50	8.85	2.02	—

*— denotes failure to reach significance at the .05 confidence level.

As indicated by Table 5, the patients who completed the test successfully were less efficient in doing so than were the comparable subjects in the control group. As a group they made more errors and required a larger number of trials to complete the test. The schizophrenic subjects also displayed a greater tendency to continue sorting on the basis of the previously correct category, as indicated by the significantly higher frequency of perseverative responses. It is interesting to note that in the records of both control and schizophrenic subjects, high concentrations of perseverative responses tended to occur either at the beginning or at the end of the different stages of the test. The patterning of these responses suggests that perseverative responses tended to occur in response to certain demands in the situation, and that the *manner* in which they were utilized by patients and controls was similar. Perseverative responses rarely constituted more than half of the total responses made by subjects in either group on any stage of the test. None of the subjects adhered consistently to this form of response on any stage of the test, and, thus, all demonstrated the ability to reject the former basis of classification at least temporarily and attempt other alternatives.

Since all subjects in these comparative groups made six runs of ten consecutive correct responses in completing the test successfully, the *CLR* measure utilized here includes only those correct responses occurring in runs of from three to nine in length. Although the tendency of the schizophrenic group to make more responses of this nature is not significant at the .05 level of confidence when corrected for heterogeneous variance, it is in the expected direction. The patient group seemed to experience more difficulty in holding the correct set, once having achieved it (in terms of this criterion), than did the control group. This tendency suggests the possible importance of motivational and attentional factors in producing the inferior performance of the schizophrenic group.

Comparison of Total Patient and Control Groups. The first seven measures in Table 6 are comparable to those in Table 5. Since, however, the groups compared include subjects failing to complete the test, the measures are based upon the means of the scores obtained in the stages completed, rather than upon total scores as in Table 5. As indicated by the score *PR in CC/CC — 1* in Table 6,[6] the higher frequency of perseverative responses in the patient group was again revealed in the

[5]There is some evidence that this may be true throughout the range. Feebleminded children have been found to be capable of successful performance on the WCST [21].

[6]The denominator, Categories Completed, is reduced by one, since by definition perseverative responses cannot occur on the first stage of the test.

TABLE 6

COMPARISON OF TOTAL PATIENT AND CONTROL GROUPS ON THE CARD SORTING TEST

Source	df	Means: Patients	Means: Controls	t	Confidence Level*
TE in CC / CC	66	6.57	5.42	0.98	—
PR in CC / CC — 1	59	5.65	3.77	2.52	.05
NPE in CC / CC	66	3.88	2.72	1.19	—
UR in CC / CC	66	0.26	0.27	0.10	—
CR in CC / CC	66	14.43	13.49	0.83	—
TT in CC / CC	66	1.08	18.87	1.00	—
CLR in CC / CC	66	2.14	1.76	0.65	—
PR / TE	66	58.05	42.30	2.95	.01
CLR / TT	66	37.82	60.69	6.02	.01

*— denotes failure to reach significance at the .05 confidence level.

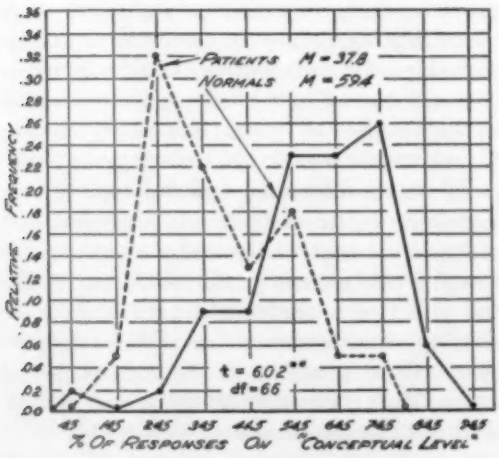

FIG. 1. Distribution of conceptual-level responses in the patient and control groups.

comparisons of the total groups, although the difference between the groups was not as marked. The significantly higher ratio of perseverative responses to total errors, the eighth measure in Table 6, further substantiates the greater tendency toward perseveration in the schizophrenic group as a whole.

The last measure in Table 6 is presented graphically in Figure 1. The *CLR* measure utilized here represents correct responses occurring in runs of three or more, including the ten consecutive correct responses which occur in successful completion of a stage of the test. Since the latter responses make up the bulk of this score for any given subject, it reflects the general degree of success achieved. Therefore, the difference between the groups indicated by this measure merely provides further confirmation of the relatively inferior performance of the patient group on the WCST as a whole. However, the distribution of this score in the two groups provides an additional basis of comparison. The score *CLR/TT* utilized in Figure 1 represents the percentage of total responses which were apparently made on the basis of at least temporary insight into the correct concept. It is interesting to note that the subjects in both groups varied widely in the extent to which they were able to utilize the correct conceptual approach, as defined by the *CLR* measure. Control subjects are found at the extremes of the range. One failed to make any responses of this nature, and three approached the maximum percentage possible on the test. This distribution strongly suggests that inefficiency in utilizing what Goldstein terms the "abstract approach" (as it may be involved in this test) occurs in presumably normal individuals in the absence of any apparent pathological process. This observation, of course, raises the possibility that we are dealing with a difference that is primarily quantitative rather than qualitative in nature, and suggests that we might further our understanding of this phenomenon through studying the conditions under which it tends to occur in normal individuals.

Comparison of Patients and Controls Failing to Complete the WCST. Table 7 presents for the two failing groups the data comparable to the first seven measures of Tables 5 and 6. In terms of these measures, it is evident that there were no differences between the groups in their performance on the stages of the test which they completed successfully. It is interesting to note, however, that the previously demonstrated tendency toward a high-

TABLE 7

COMPARISON OF PATIENTS AND CONTROLS FAILING TO COMPLETE THE CARD SORTING TEST

(BASED ON SCORES IN THE STAGES COMPLETED)

Source	df	Means Patients	Means Controls	t	Confidence Level*
TE in CC	21	6.25	8.01	0.55	—
CC					
PR in CC	14	5.21	7.08	0.88	—
CC — 1					
NPE in CC	21	4.20	5.27	0.39	—
CC					
UR in CC	21	0.32	0.61	0.63	—
CC					
CR in CC	21	14.14	16.30	0.80	—
CC					
TT in CC	21	20.39	24.34	0.68	—
CC					
CLR in CC	21	1.85	3.39	1.13	—
CC					

*— denotes failure to reach significance at the .05 confidence level.

er frequency of perseverative responses in the patient group is no longer present. In fact, the tendency is reversed, although the difference between the groups is not a significant one. It was also noted that the mean frequency of perseverative responses in the failing schizophrenic group was lower than that in the passing schizophrenic group. These relationships suggest that perseveration may be characteristic of an intermediate level of performance on this test. In a sense, the subject who makes a large number of perseverative responses demonstrates greater ability to maintain set, albeit an incorrect one, than does the subject whose performance is characterized by an apparently random choice of responses.

TABLE 8

COMPARISON OF PATIENTS AND CONTROLS FAILING TO COMPLETE THE CARD SORTING TEST

(BASED ON THE SCORES ON THE FAILED STAGE)

Source	df	Means Patients	Means Controls	t	Confidence Level*
CR	22	40.37	27.62	1.07	—
PR	22	23.50	14.50	1.72	—
NPE	22	19.56	22.50	0.69	—
UR	22	3.12	1.62	0.69	—
CLR	22	9.44	12.63	0.66	—
TE	22	23.63	36.68	1.07	—

*— denotes failure to reach significance at the .05 confidence level.

The data of Table 8 are based upon total scores on the stages of the test at which failure occured. In terms of the measures utilized, it is evident that there were no significant differences between the control and schizophrenic groups in their performance on the failed stage of the test. However, further analysis revealed an interesting difference between the groups. Inspection of the data revealed a marked tendency on the part of the subjects in the control group to make several runs of three or more consecutive correct responses (*CLR*) in their performance on the stage of the test which they were unable to complete successfully. This tendency was analyzed in terms of the numbers of subjects in each of these groups making and failing to make two or more such runs on the failed stage of the test. Only 1 of the 8 control subjects failed to meet this criterion, while 11 of the 16 patients failed to do so. A chi-square test of these data, utilizing the Yates correction for continuity, revealed the difference to be significant at the .05 level of confidence. In other words, there was a greater tendency on the part of the subjects in the control group to make responses on the failed stage of the test which are considered to be indicative of partial or transitory insight into correct solution. The tendency toward reversal in the direction of the *CLR* measure in comparing the successful and failing groups suggests that it reflects an unstable type of performance, intermediate to highly efficient success and complete failure. In terms of this measure, the patient group which passed the test exhibited a tendency toward greater instability of performance than did the comparable control group, and the control group which failed the test evidenced greater insight than did the comparable patient group.

TABLE 9

NUMBER OF PATIENTS AND CONTROLS FAILING AT EACH STAGE OF THE TEST

Group	(1) Color	(2) Form	(3) Number	(4) Color	(5) Form	(6) Number	Total
Patients	0	4	7	5	0	0	16
Controls	1	3	0	3	1	0	8

Analysis of the stages of the test at which failure occurred, presented in Table 9, suggests further differences between the two failing groups. The most striking difference occurred in response to the number category on the first half of the test. Forty-three per cent of the failures in the patient group occurred at this stage of the test, while none of the control subjects who reached this stage of the test failed to pass it. Heidbreder's work with concept formation [13, 14, 15, 16, 17, 18, 19, 20] indicates that number concepts are formulated with greater difficulty than are color and form concepts. This conclusion is supported by the performance of the present schizophrenic group. However, an apparent contradiction is evident in the performance of the control group. The results of this and other studies [10] with the WCST suggest that when it appears in this position in the test, the number category is more easily solved by normal subjects than are the color and form categories. Recent research on the WCST suggests an explanation of this apparent discrepancy. Grant [8] demonstrated that college students apparently utilize positional cues to a large extent in solving the number category. These subjects experienced markedly greater difficulty in solving the number category when positional cues were eliminated, and they were thus forced to utilize an analytical or numerical approach to solution. It is suggested, therefore, that the marked difference between the control and schizophrenic groups in their performance on the number category may be indicative of an inability on the part of the schizophrenic subjects to utilize a perceptual or configurational approach and respond to the pattern as a whole rather than to the stimulus properties of the individual figures. It should be noted here that, unlike the number category, the color and form categories can be successfully solved by responding to the stimulus properties of the individual figures on the cards. In general then, the fact that the schizophrenic subjects experienced greater difficulty in solving the number category than the color and form categories may be related to the tendency of these patients to respond to the more detailed aspects of the stimulus situation.

The relatively high incidence of failure in both groups on the fourth stage of the test is interesting, since all of these subjects had previously demonstrated the ability to deal effectively with the color concept on the first stage of the test. On the basis of verbal reports, our best guess concerning the source of these failures was a certain rigidity in interpretation of the examiner's judgments, a "once wrong, always wrong" approach. Since color matches had been judged "wrong" on the preceding two stages of the test, several of these subjects expressed the feeling that it could no longer be considered as a possibility. Analysis of these failures serves to emphasize the importance of at least two factors in determining the nature of an individual's response to situations of this kind: the subject's own interpretation of the task, and the actual structure of the situation in which his responses are elicited. In general, it serves as a caution against interpreting failure to use a given concept in a particular situation as sufficient evidence of inability to utilize it in general.

Verbal Reports. Analysis of verbal reports made by the subjects at the conclusion of the WCST revealed more similarities than differences between the control and patient groups. In considering these observations, it should be remembered that the patient group was, on the whole, well organized and capable of effective communication.

Some of the subjects in both groups utilized concrete or specific terms rather than general category names in describing their performance. A typical response of this nature would be, "I put the triangles and crosses and circles together" rather than, "I matched them by shape or form." The number of subjects utilizing such responses in referring to each of the major categories is indicated in Table 10. In general, the relative frequency of this

TABLE 10

NUMBER OF SUBJECTS USING CONCRETE TERMS IN DESCRIBING MAJOR CATEGORIES

	Color	Form	Number
Patients	1	8	5*
Controls	1	6	11

*Only 10 of the patients made any attempt to describe this category.

type of response in the two groups is similar. The same order of difficulty is also indicated. It is interesting to note that the control subjects also experienced the greatest difficulty in verbalizing the number category in categorical terms, although none of them failed at this stage of the test.

The verbal reports of both control and schizophrenic subjects were also found to show a high degree of correspondence to actual performance. Lack of correspondence occurred in only three reports, one by a control subject and two by patients. This is in contrast to the frequent lack of correspondence noted by Cameron [6] in the explanations offered by his schizophrenic subjects of their performance on the Vigotsky test. Both the relative absence of serious disorganization in the patients of the present group and the relative simplicity of the WCST were probably factors contributing to this discrepancy.

Although no systematic attempt was made to evaluate emotional tension during the testing procedures, it was the impression of the examiner that severe tension may have a disrupting effect upon performance on the WCST. Two of the subjects in the control group who displayed the most marked and obvious signs of tension throughout testing failed to complete the sorting test. One was described as "extremely nervous" by his foreman, and the other repeatedly verbalized his anxiety by such comments as, "This is really getting me down."

SUMMARY AND CONCLUSIONS

The Wisconsin Card Sorting Test and the Wechsler-Bellevue Adult Intelligence Scale were administered to a group of 22 young schizophrenic patients and 47 young normal subjects. The groups were matched in age, intelligence, educational level, and ratio of men to women. In terms of gross measures of success on the WCST, the schizophrenic group was found to be clearly inferior to the control group. No significant correlations were found between Wechsler Full Scale IQ or subtest scores and success on the WCST.

Comparison of the performance of the schizophrenic and control subjects who passed the test revealed the following differences: (1) higher total trial and total error scores in the schizophrenic group, (2) higher frequency of perseverative responses in the schizophrenic group, and (3) a tendency toward a higher frequency of correct responses in runs of from three to nine in length in the schizophrenic group.

Comparison of the performance of the total schizophrenic and control groups on the WCST revealed the following relationships: (1) higher frequency of perseverative responses in the schizophrenic group, and (2) wide variability in the percentage of conceptual-level responses made by the subjects in both groups.

Comparison of the performance of the schizophrenic and control subjects failing to complete the WCST revealed the following differences: (1) a greater tendency toward the occurrence of conceptual-level responses on the failed stage of the test in the control group, and (2) relatively greater incidence of failure on the number category at the third stage of the test in the schizophrenic group.

Analysis of the verbal reports of performance on the WCST revealed more similarities than differences between the groups.

In general, the major results of this study support the findings of previous studies utilizing similar testing techniques with schizophrenic patients. The schizophrenic subjects in this study experienced greater difficulty in solving a problem generally assumed to involve the effective utilization of generalizations or concepts than did the normal group. Within the structure of this study it can be stated that this discrepancy cannot be attributed to differences in age, intelligence, or educational level. And in view of the nature of the patient group utilized, it must be concluded that the factors responsible for the observed deficiency in performance on the WCST are present in the early stages of schizophrenic illness, and in the absence of evidence of serious disorganization or the deteriorative effects of prolonged illness.

More important than the demonstration of gross differences between these groups, it is hoped that the detailed analysis of performance made possible by the testing technique utilized has served to point up some of the factors

which may prove to be of importance in accounting for these differences. Although such aspects of performance as perseveration, difficulty in maintaining set, and difficulty in responding to patterns as wholes might all be subsumed under some such explanatory concept as "loss of abstract attitude," it is doubtful that they all emanate from a single psychological "cause," or that our understanding of schizophrenic behavior will be furthered by making such assumptions. It would seem that what is needed, rather, is an attempt to isolate and measure characteristic aspects of behavior and to study the conditions under which they tend to occur in normal as well as abnormal individuals. The importance of research with normal groups in this area is particularly emphasized by the finding in this study that none of the behavior found to be characteristic of the schizophrenic subjects was limited to that group, but always occurred to varying degrees in the control group.

Received December 18, 1950.

References

1. Berg, Esta A. A simple objective technique for measuring flexibility in thinking. *J. gen. Psychol.*, 1948, **39**, 15–22.
2. Bolles, M. Marjorie, and Goldstein, K. A study of impairment of "abstract" behavior in schizophrenic patients. *Psychiat. Quart.*, 1938, **12**, 42–65.
3. Cameron, N. Reasoning, regression and communication in schizophrenics. *Psychol. Monogr.*, 1938, **50**, No. 1.
4. Cameron, N. A study of thinking in senile deterioration and schizophrenic disorganization. *Amer. J. Psychol.*, 1938, **51**, 650–665.
5. Cameron, N. Deterioration and regression in schizophrenic thinking. *J. abnorm. soc. Psychol.*, 1939, **34**, 265–270.
6. Cameron, N. Schizophrenic thinking in a problem-solving situation. *J. ment. Sci.*, 1939, **85**, 1012–1035.
7. Goldstein, K., and Scheerer, M. Abstract and concrete behavior: An experimental study with special tests. *Psychol. Monogr.*, 1941, **53**, No. 2.
8. Grant, D. A. Perceptual versus analytic responses to the number concept of a Weigl-type card sorting test. *J. exp. Psychol.*, 1951, **41**, 23–29.
9. Grant, D. A., and Berg, Esta A. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *J. exp. Psychol.*, 1948, **38**, 404–411.
10. Grant, D. A., Jones, O. R., and Tallantis, Billie. The relative difficulty of the number, form and color concepts of a Weigl-type thinking and reasoning problem. *J. exp. Psychol.*, 1949, **39**, 552–557.
11. Hanfmann, Eugenia, and Kasanin, J. A method for the study of concept formation. *J. Psychol.*, 1936, **3**, 521–540.
12. Hanfmann, Eugenia, and Kasanin, J. Conceptual thinking in schizophrenia. *Nerv. ment. Dis. Monogr.*, 1942, No. 67.
13. Heidbreder, Edna. Toward a dynamic theory of cognition. *Psychol. Rev.*, 1945, **52**, 1–22.
14. Heidbreder, Edna. The attainment of concepts—a psychological interpretation. *Trans. N. Y. Acad. Sci., II*, 1945, **7**, 171–188.
15. Heidbreder, Edna. The attainment of concepts: I. Terminology and methodology. *J. gen. Psychol.*, 1945, **35**, 173–189.
16. Heidbreder, Edna. The attainment of concepts: II. The problem. *J. gen. Psychol.*, 1946, **36**, 191–223.
17. Heidbreder, Edna. The attainment of concepts: III. The process. *J. gen. Psychol.*, 1947, **24**, 93–138.
18. Heidbreder, Edna. The attainment of concepts: IV. Regularities and levels. *J. gen. Psychol.*, 1948, **25**, 299–329.
19. Heidbreder, Edna, and Overstreet, Phoebe. The attainment of concepts: V. Critical features and contexts. *J. Psychol.*, 1948, **26**, 45–69.
20. Heidbreder, Edna. The attainment of concepts: VI. Exploratory experiments on conceptualization at perceptual levels. *J. Psychol.*, 1948, **26**, 193–216.
21. Magaret, Ann, Grant, D. A., and Berg, Esta A. A study of the performance of exogenous and endogenous feebleminded children on the Wisconsin Card Sorting Test. (Unpublished.)
22. Vigotsky, L. Thought in schizophrenia. *Arch. Neurol. Psychiat.*, Chicago., 1934, **31**, 1063–1077.
23. Wegrocki, H. Generalizing ability in schizophrenia. *Arch. Psychol., N. Y.*, 1940, No. 254.
24. Weigl, E. On the psychology of the so-called processes of abstraction. *J. abnorm. soc. Psychol.*, 1941, **36**, 3–33.

NERVOUS SYMPTOMS AMONG MILITARY PERSONNEL AS RELATED TO AGE, COMBAT EXPERIENCE, AND MARITAL STATUS[1]

RAYMOND G. KUHLEN
SYRACUSE UNIVERSITY

THE military situation in which the present data were collected offered unusual opportunity to study adult age trends in adjustment. In common with other military personnel, the subjects of this investigation had been removed by circumstances of war from accustomed civilian environments and forced into new situations of varying degrees of stress and trauma—but in addition, the present subjects were, at the time of the study, in a somewhat stressful situation. How did they react? Do older men within the military age range show more tension and difficulties of adjustment than younger men? Do the age trends in nervous symptoms among those who have been subjected to the stresses and strains of combat or to the tensions of overseas noncombat experience differ from age trends apparent among military men who have not yet been out of the boundaries of the country? Do married men with family responsibilities manifest more nervous symptoms after being exposed to combat than their unmarried contemporaries? The data presented in this paper are relevant to these questions.

PROCEDURE

The subjects of this investigation were some 8,700 enlisted naval personnel, representing a portion of those who passed through a preembarkation center during a six-month period from October, 1944, through March, 1945. Within 72 hours after entering the center, they would embark for unknown destinations in the Pacific combat zone. Standard processing of the men at this point included psychologic screening in an effort to eliminate those who seemed poor psychological risks for overseas duty. Since it was impossible to interview each man, the procedure involved the administration of a brief neurotic inventory (described below) to the total group, with subsequent interview and special further testing where desirable, of those who indicated nervous symptoms in excess of a particular cutting-score. At the time of taking this inventory each subject provided certain information about himself, including such facts as his age, marital and family status, and the nature of his previous military experience—in particular whether his military duty had included combat experience (actually under fire), overseas duty without combat experience, or only duty within the boundaries of the United States. The 8,702 cases upon which this report is based were distributed among these three groups in this way: 2,038 had overseas experience with combat; 1,521 had overseas experience but without combat; and 5,143 had military experience limited to the United States. It is to be noted that not only did the subjects differ in the extent to which they had been subjected to stress thus far in their military careers, but that they were at the time of testing in what was essentially a stress situation: they were scheduled for departure to unknown combat destinations at any minute within a maximum of 72 hours. Table 1 contains an analysis of the sample studied, with respect to the three ex-

[1]Read at the 1949 meeting of the American Psychological Association in Denver, Colorado. The data presented here were collected while the writer was on active duty in the Navy. The opinions expressed are, of course, those of the writer and are not to be construed as reflecting the official attitude of the Navy Department.

TABLE 1
PERCENTAGE DISTRIBUTION ACCORDING TO AGE OF THE 8,702 MEN STUDIED, GROUPED ACCORDING TO COMBAT EXPERIENCE

Age	U.S. Only		Overseas Noncombat		Overseas Combat	
	N	%	N	%	N	%
18–20	1,513	29.4	385	25.3	522	25.6
21–23	884	17.2	416	27.3	668	32.8
24–26	844	16.4	220	14.5	360	17.7
27–29	586	11.4	168	11.0	171	8.4
30–32	530	10.3	128	8.4	131	6.4
33–35	405	7.9	93	6.1	79	3.9
36–38	296	5.8	70	4.6	56	2.7
39+	88	1.7	42	2.8	51	2.5
Totals	5,143	100.1	1,521	100.0	2,038	100.0

perience groups and various age categories.

Information regarding "nervous symptoms" was obtained by means of the "Personal Inventory," a brief 20-item forced-choice inventory which had been designed specifically for military screening purposes. The inventory is made up of the 20 most discriminating items from a longer 145-item inventory of the same name; the two forms correlate .84 [2]. Reports regarding the reliability of this short inventory indicate Spearman-Brown corrected split-half correlations based on various groups, of the magnitude of .66, .68, .89, .87, .91 [2]. Validity is implied in the finding that when tests were administered to unselected groups who were later screened by interview for psychiatric disabilities without examiners' having prior knowledge of the test score, particular cutting-scores would have resulted in the identification of 52.5 per cent of those discharged for psychiatric reasons while including less than 7 per cent of those who were passed by psychiatric interview. Further, the inventory discriminated reliably between patients on psychiatric wards and unhospitalized men [2]. Studies bearing on the validity of this and other "personality" inventories in military practice have been reviewed critically by Ellis and Conrad [1].

One may well question the validity of responses obtained at pre-embarkation when there exists the possibility that a man would be removed from the "draft" if he were notably "nervous." This possibility might conceivably result in a general biasing of responses, particularly on the part of those who had already seen combat or had been overseas and who had no desire to go again. While no direct evidence is available on this question, the writer's clinical impression of the men studied was in agreement with the findings presented below.

FINDINGS

The results of this testing have been analyzed in several ways. In Table 2 the findings are presented in terms of the proportion of the various groups who reached or exceeded a particular cutting-score and thus were subjected to further interview study. It is immediately clear that those with combat experience and those in the older age groups contributed disproportionately large numbers to the group interviewed. While 6.7 per cent of those with only United States experience reached or exceeded a cutting-score of 12, 12.6 per cent of those with overseas noncombat experience and 17.6 per cent of those with combat experience reached or exceeded this cutting-score. These percentages differ reliably. In contrasting younger and older age groups, it will be noted that for the two oldest age groups a reversal in trend occurs among those with overseas experience, and it will be noted also on inspection of Table 1 that the numbers of cases in the oldest groups are relatively small. These two facts, especially the reversal in trend, suggest that selective factors (which one might expect to operate at the upper ex-

TABLE 2
PERCENTAGE OF DIFFERENT AGE GROUPS WITH VARYING NAVAL EXPERIENCE WHO CHECKED A LARGE NUMBER (12 OR MORE OUT OF 20) OF NERVOUS SYMPTOMS

Age	U.S. Only	Overseas Noncombat	Overseas Combat
18–20	3.1	6.2	14.8
21–23	5.2	8.9	15.6
24–26	7.0	13.2	17.2
27–29	8.2	14.3	21.1
30–32	10.4	17.2	29.8
33–35	10.1	35.5	29.1
36–38	14.2	25.7	14.3
39 over	10.5	11.9	17.6
Total Group	6.7	12.6	17.6

treme and beyond the military draft age) have resulted in sampling differences which make it reasonable to eliminate these groups from consideration of age differences. It will be noted that for all three experience groups there is a fairly regular increase from ages 18 to 20 to ages 33 to 35 in percentage exceeding the cutting-score.

TABLE 3

MEAN NUMBER OF NERVOUS SYMPTOMS FOR EACH AGE GROUP FOR THE THREE EXPERIENCE GROUPS, AND CR'S SHOWING RELIABILITY OF DIFFERENCES AMONG THE EXPERIENCE GROUPS OF VARIOUS AGES

	Age Groups					
	18–20	21–23	24–26	27–29	30–32	33–35
Mean Scores						
U.S. Only	3.2	3.9	4.5	5.0	5.3	5.3
Overseas Noncombat	4.8	5.0	5.8	6.1	6.9	8.2
Overseas Combat	6.0	6.2	6.6	7.2	8.2	8.7
Critical Ratios						
Between U.S. Only and Overseas Noncombat	8.1	4.4	3.8	3.0	3.7	5.2
Between Overseas Noncombat and Overseas Combat	4.1	4.5	2.0	2.2	2.2	0.8
Between U.S. Only and Overseas Combat	13.1	10.6	8.0	5.8	6.0	7.1

While these percentages indicate that age and combat experience bear a relation to number of nervous symptoms, the main analysis of the data consists of a study of the significance of differences among means. A preliminary analysis of variance had already yielded *F*'s substantially larger than those required for confidence levels of .01, indicating that both age and "experience" were reliably related to number of nervous symptoms. Table 3 contains the mean number of nervous symptoms for each age and experience category (the top half of the table) and the *CR*'s (the lower half of the table) showing the reliability of differences among the experience groups of comparable ages. For the reasons outlined above, the data for the two highest age groups (36–39 and 39 plus) have been eliminated from further consideration. The means of the three experience groups differ at all ages, with the groups having United States experience only being consistently the lowest in number of symptoms expressed, and the overseas combat group being consistently the highest. The *CR*'s indicate that these differences are in most comparisons highly reliable. Only in the contrasts between the overseas noncombat and the overseas combat groups do the *CR*'s shrink to the area between the 1 and 5 per cent levels of confidence, and in these contrasts the *N*'s were relatively small. In only one instance does the *CR* fall below that required for the 5 per cent level of confidence.

TABLE 4

CONTRASTS BETWEEN SELECTED AGE GROUPS WITH RESPECT TO INCREASE IN NUMBER OF NERVOUS SYMPTOMS FOR THE THREE EXPERIENCE GROUPS

	Age			
	27–29	33–35	Diff.	$\sigma_{Diff.}$
(A) U. S. Only	4.99	5.26	0.27 (A)*	.2584
(B) Overseas Noncombat	6.14	8.18	2.04 (B) *	.6205
(C) Overseas Combat	7.22	8.73	1.51 (C) *	.5704

*The difference between differences "A" and "B" is 2.63 times its standard error (level of confidence, .01); the difference between differences "A" and "C" is 1.99 times its standard error (level of confidence, .05).

Inspection of these means reveals that in the upper age groups represented in Table 3, the number of nervous symptoms increases for the overseas groups (both combat and noncombat) but tends to level off for those with only United States experience. Are these differences in gains over this older portion of the age range reliable? Table 4 focuses attention on the appropriate contrasts. The United States age groups of 27 to 29 and 33 to 35 differed by 0.27 in mean number of nervous symptoms; the overseas noncombat group in these same age categories differed by 2.04 in mean number of symptoms; and the same age contrasts for the combat groups revealed a difference of 1.51 between the means. The reliability of the difference between differences is demonstrated by the *CR*'s noted in the footnote to Table 4. One is at the 5 per cent level of confidence; the other at the 1 per cent level. It would appear, then, that during this

age range, age increases in nervous symptoms are apt to appear among those previously exposed to substantial war stress, whereas among those who have not been previously subjected to such stress, such age trends are not so apparent. This finding has certain methodological as well as practical implications, as will be pointed out later.

TABLE 5
MEAN NUMBER OF NERVOUS SYMPTOMS CHECKED BY INDIVIDUALS WHO WERE SINGLE, MARRIED WITH NO CHILDREN, OR MARRIED WITH CHILDREN

	Age Group					
	18–20		21–23		24–26	
	Means	N	Means	N	Means	N
Means						
Unmarried	5.7	445	5.5	443	5.8	185
Married, no children	7.6	56	8.2	156	7.2	100
Married, with children	8.1	20	6.4	38	7.6	67
Critical Ratios						
Between Unmarried and Married, no children	3.3		6.0		2.5	
Between Unmarried and Married, with children	2.2		1.9		3.2	
Between Married, no children, and Married, with children	0.4		1.1		0.7	

What is the relationship of marital status to nervous symptoms evidenced in the manner and under the conditions described in this study? A change of duty stations on the part of the writer prevented a complete analysis of the relevant data, but certain facts were summarized for the younger age groups among those who had been subjected to combat. The appropriate means and *CR*'s are presented in Table 5. It will be noted that those who were married, whether or not they had children, expressed reliably more nervous symptoms on the questionnaire than did the unmarried. The differences between married individuals with and without children were not reliable. Apparently, the existence of marital and family responsibilities results in a greater reaction to the stress of war experiences and the expression of more "nervous" symptoms. Ordinarily, married individuals as a group might be expected to be better adjusted than the unmarried.

CONCLUSIONS AND DISCUSSION

Although the conditions of the study (the military situation, the possible biasing of response, and the limitations of the instrument) naturally impose limitations upon interpretations, three conclusions are suggested. In the first place, the data suggest that important age trends in psychological adjustment occur during the early adult years, and, thus, these years should not be neglected by those who wish to understand the aging process. The current tendency for gerontological research to be focused upon old age alone represents an undue and unwise limitation of the field.

In the second place, it would appear from the data here presented that studies of the relation of adult age to psychological adjustment should utilize stress situations. It will be recalled, for example, that in the present data the "effect" of aging upon adjustment, that is, upon frequency of nervous symptoms, was more evident among those who had been previously subjected to stress than among those who had not been so subjected. But it is noteworthy that age trends, particularly in the younger years studied, were present even among those who had only United States experience, whereas another investigation [3] of recruits had revealed no correlation between age at entrance to the Navy and score on the long form of this inventory. In the present study the occasion of testing (pre-embarkation) was in itself probably somewhat stressful (thus accounting for the age trends in the "U.S. only" group), whereas in the study just cited such stress presumably was lacking. Physiologists have long been aware that the effects of aging on physiological functions may not be evident under normal circumstances but yet be clearly apparent when the organ is subjected to stressful loads. Psychologists interested in aging have not yet utilized such methodology.

And, finally, there is a suggestion for further research leading to practical applications. In the military situation, further study would seem warranted to determine (*a*) the optimal

ages of personnel for specific stressful tasks and (*b*) the extent to which selective placement (e.g., service limited to the United States) would permit the use of older selectees. Both are important considerations in the event of total war mobilization of human resources. In civilian life, there seems to be some indication that younger men are better able to "take" the competition of business and industrial life and that some older individuals, if they are to maintain their good adjustment, might advisably shift into less stressful though perhaps no less responsible positions. Though occupational and other stresses of civilian life are less subject to precise definition and experimental control (compared with the military situation), general studies of capacity of older age groups to resist stress would provide pertinent information.

Received November 1, 1950.

References

1. Ellis, A., and Conrad, H. S. The validity of personality inventories in military practice. *Psychol. Bull.*, 1948, **45**, 385–426.
2. Shipley, W. C., and Graham, C. H. Final report in summary of research in the Personal Inventory and other tests. (OSRD, 1944; Publ. Bd., No. 12060) Washington: U. S. Dept. Commerce, 1946.
3. Shipley, W. C., Gray, R. W., and Newbert, N. A comparison of Personal Inventory scores with service records one year after testing. (OSRD, 1944; Publ. Bd., No. 12061) Washington: U. S. Dept. Commerce, 1946.

THE WECHSLER-BELLEVUE AND DESCRIPTIVE DIAGNOSIS

J. R. WITTENBORN[1]
YALE UNIVERSITY
AND
J. D. HOLZBERG
CONNECTICUT STATE HOSPITAL

THERE are two respects in which the Wechsler-Bellevue test of intelligence has been used as an aid in psychiatric diagnoses. Possibly the older and more valid diagnostic use has been in attempts to distinguish patients whose behavorial disturbance is due to brain pathology from patients whose behavioral disturbance may be considered to be functional. This distinction, aimed at etiological factors, is based on the presumption that the thought processes of patients with certain forms of brain pathology are different from the thought processes of many other patients. The present report does not examine this use of the Wechsler-Bellevue.

A second diagnostic use of the test is related to a general belief that there are systematic differences in thinking among patients whose behavior disturbance is considered to be functional. Presumably, such differences in thinking are expressed in the patient's performance on a complex test of intelligence such as the Wechsler-Bellevue, and presumably these differences affect the patient's diagnosis. Accordingly, it has been proposed that characteristics of a patient's performance on the Wechsler-Bellevue may be usefully employed as an aid in determining which functional diagnostic group the patient most resembles. Unfortunately, the value of the Wechsler-Bellevue test in making distinctions between patients who may be placed in different diagnostic groups is not clear. An ambitious proposal to use the test in this way [3] has been followed by discouraging empirical tests [1, 2, 6].

The various attempts to use the test as a diagnostic aid have been based upon an employment of intra-individual differences in subtest scores. Many investigators have shown an almost inexplicable lack of concern for the fact that there are marked intra-individual differences of various kinds for individuals without known mental pathology. As has been shown for certain well-known cases, the variety of intra-individual test score differences within a diagnostic group is of about the same order as the variety of intra-individual differences between diagnostic groups [6]. There is also indication that some important investigators in this field have been inclined to believe that differences between subtest scores such as revealed by profile analysis may be employed to predict diagnostic criteria which could not be predicted from a direct use of the various subtest scores. This latter point of view has been examined in a recent discussion and appears to be indefensible [8].

The present report describes an attempt to evaluate the diagnostic validity of the Wechsler-Bellevue test by an employment of two different types of criteria. The sample comprises patients at the Connecticut State Hospital at Middletown who had been tested by the Wechsler-Bellevue during an arbitrarily selected period of time and were without known neural defects. The sample is considered to be similar to others wherein the Wechsler might typically be used as a diagnostic aid.

[1]The data for the present study were provided by the courtesy of the Staff of the Connecticut State Hospital at Middletown. The support of the Institute of Human Relations, Yale University, and the United States Public Health Service is gratefully acknowledged.

Because of the nature of the sample, the results of the present analysis may not safely be generalized to other samples such as college students or individuals seen in outpatient clinics.

The nature of the criteria and the respective analytical procedures employed in the present report follow:

1. Psychiatrists' diagnoses. From the available sample, the four most commonly occurring diagnostic groupings were selected: paranoid schizophrenia, manic-depressive manic phase, alcoholic psychosis,[2] and psychopathic personality. Because earlier attempts to show systematic intertest profile differences between the various diagnostic groups were discouraging [6], the present data are analyzed differently. Differences between the various diagnostic groups are sought for each of the eleven subtests of the Wechsler-Bellevue. In this way evidence for correlation between subtest score and membership in one of the four diagnostic groups may be revealed. Inasmuch as a psychiatric diagnosis is not a quantified attribute, the presence of relationship is sought through an application of the χ^2 test. This is illustrated in Table 1, which shows that the alcoholic patients tend to do less well than expected on the similarities test while the psychopathic patients do better than is expected for this subtest. In Table 1, as in all of the χ^2 tests, the expected performance is based on the distribution of scores for a group of 82 patients, resulting when the patients in the four diagnostic groups are combined.

TABLE 1

DISTRIBUTIONS OF SIMILARITY SUBTEST SCORES FOR FOUR DIFFERENT DIAGNOSTIC GROUPS

(Distributions of subtest scores predicted by an hypothesis of no relationship are given in the parenthesis.)

	Paranoid Schiz.		Manic-Depressive Manic		Alcoholic Psychosis		Psychopathic Personality		
Score on Similarities Subtest	Observed Frequency	Expected Frequency	Observed Frequency	Expected Frequency	Observed Frequency	Expected Frequency	Observed Frequency	Expected Frequency	Sum
0–8	8	(7)	6	(6)	16	(11)	2	(8)	32
9–12	4	(5)	4	(4)	7	(6)	5	(5)	20
13–19	7	(7)	5	(5)	4	(10)	14	(8)	30
Σ	19		15		27		21		82

2. Syndrome scores. The criteria for the second analysis are the patients' scores for each of nine different empirically determined symptom clusters [7]. The patient's score on each of these nine clusters is based on psychiatrists' ratings and provides a kind of a quantified multiple psychiatric diagnosis. The nine scores evaluate the patient in terms of most of the different aspects of behavior which are currently considered by psychiatrists to be important in the description of mental hospital patients [7]. This type of criterion should be more favorable for revealing a relationship between Wechsler subtest scores and kind of behavior disturbance than psychiatric diagnoses, per se, as criteria. There is an unknown degree of overlap in psychiatric diagnoses, and for this reason a diagnosis is an unprecise criterion of uncertain behavioral significance whereas the cluster scores are based on separate groups of rated symptoms; an additional favorable feature of the second criterion is its quantified nature; its use can show the degree of relationship between subtest scores and behavioral disturbances.

RESULTS

The results of Analysis 1 cannot be taken as evidence for a practically useful degree of relationship between Wechsler-Bellevue subtest scores and the four diagnostic groups employed in the present sample. Scores for the information, digit span, arithmetic, vocabulary, block design, and object assembly subtests are uniformly distributed for each of the four diagnostic groups. Accordingly, it seems improbable that these six subtests of the Wechsler-Bellevue may be profitably employed as an aid in the differential diagnosis of the four different diagnostic groups. There is a significant tendency for the psychopathic patients to be superior to the other patients in the comprehension, similarities, picture arrangement, picture completion, and digit symbol substitution subtests. Inasmuch as the diagnostic term, "psychopath," is usually ap-

[2]Since alcoholic psychosis is a toxic psychosis, this diagnosis is an exception to the functional classification with which the report is primarily concerned.

plied to patients suffering from a character disorder and without affective or cognitive disturbances which detract from ordinary thought processes, this is not an untoward finding. In a sample of mental hospital patients, it is reasonable to find that patients whose thought processes are relatively unaffected by their illness should also comprise a group which does relatively well on mental tests.

It is interesting to observe that the alcoholic patients do less well than would be expected on the similarities subtest. From this it may be tentatively suspected that patients hospitalized because of an alcoholic psychosis are less sensitive than other patients to differences and similarities and may suffer some impairment of discrimination.

It is concluded that there are no useful differences in Wechsler-Bellevue subtest performances among patients who are diagnosed as paranoid schizophrenia, manic-depressive manic, or alcoholic psychosis. It is not concluded that groups of patients bearing these diagnostic labels cannot be formed in a fashion that would reveal a difference in performance on the Wechsler-Bellevue subtests. It is merely concluded that (for the present diagnostic groups) mental hospital patients for whom a Wechsler-Bellevue test is likely to be employed may not differ from each other in Wechsler-Bellevue subtest performance to any practical degree.

Analysis 2 shows evidence of modest relationships between various syndrome scores and various subtests of the Wechsler-Bellevue. The analysis is based on 79 patients with functional disorders, and correlation coefficients smaller than .24 are arbitrarily rejected as insignificant. (Incidentally, the samples employed in the two analyses overlapped considerably.) The correlations are given in Table 2. The most conspicuous trend among these correlations is the general tendency for relatively high scores on the hysteria scale to be related to relatively high scores on the various subtests of the Wechsler. This tendency is reminiscent of the positive relations

TABLE 2

CORRELATIONS BETWEEN WECHSLER-BELLEVUE SUBTEST SCORES AND SYMPTOM CLUSTER SCORES FOR 79 MENTAL HOSPITAL PATIENTS

	Information	Comprehension	Digit Span	Arithmetic	Similarities	Vocabulary	Picture Arrangement	Picture Completion	Block Design	Object Assembly	Digit Symbol
Manic	—.001	—.045	—.115	—.024	—.210	—.021	—.106	.043	—.166	—.072	—.100
Excitement	—.086	—.183	—.252	—.186	—.263	—.133	—.245	—.111	—.268	—.215	—.281
Anxiety	—.037	.067	.077	—.136	.099	.042	—.023	.032	—.037	—.016	—.009
Depressed	—.151	—.179	—.123	—.284	—.070	—.079	—.181	—.173	—.182	—.177	—.156
Paranoid Schizophrenic	—.143	—.213	—.110	—.150	—.155	—.136	—.219	—.185	—.192	—.094	—.251
Paranoid Condition	—.115	—.168	—.099	—.082	—.104	—.158	—.227	—.116	—.180	—.122	—.288
Hysteria	.303	.285	.421	.130	.174	.398	.136	.343	.228	.071	.238
Phobic-Compulsive	—.013	—.021	—.050	—.142	—.017	—.003	—.103	—.054	—.127	—.156	—.115
Hebephrenic or Deteriorated Schizophrenia	—.123	—.133	—.235	—.195	—.303	—.149	—.206	—.088	—.245	—.210	—.147

between the diagnosis of psychopathic personality and relatively high scores for Wechsler subtests observed in five of the eleven comparisons provided by Analysis 1. The positive tendencies for the psychopathic diagnosis in Analysis 1 and hysterical syndrome in group 2 may be explained in similar terms. In a group of patients in a mental hospital, those with a diagnosis of psychopathic personality are likely to have their thought processes little disturbed by their symptoms. Similarly, mental hospital patients with marked hysterical manifestations are likely not to be the patients with symptoms which detract greatly from their thought processes; in many cases hysterical patients may be in the hospital primarily because of disabling physical symptoms. It should not be concluded from these data that there is a general positive relationship between mental ability and hysterical tendencies which applies to members of the general population. This positive relationship is probably found only in mental hospital samples and very likely is due to the thought disturbance suffered by most mental hospital patients (especially those who do not have conversion symptoms).

Table 2 also suggests that there is a general tendency for patients who have high scores on the excitement syndrome to make relatively low scores on a number of the subtests. A similar loss of mental ability is indicated for patients who have high scores on the hebephrenic (or deteriorated) syndrome. It is interesting to note that the only syndrome score which is associated with loss on the arithmetic subtest is the depressed syndrome. This very likely is due to the fact that the items of the arithmetic test have a time limit and that severely depressed patients are often greatly delayed in their verbal responses.

These correlations may be employed to indicate roughly the degree to which various types of symptom manifestations are accompanied by a loss of mental ability. Relative to the patients who have marked psychotic syndromes, patients with anxiety symptoms or phobic-compulsive symptoms may suffer very little impairment of mental function. Among the psychotic syndromes the patients with high manic scores appear to suffer the least mental impairment. Nevertheless, there is no evidence of conspicuous or unique losses which offer the possibility that Wechsler subtest scores may be reliably employed to indicate the symptom pattern most likely to characterize patients.

DISCUSSION

As a result of the foregoing analyses little evidence can be offered that patients who are diagnosed differently or who present different types of symptomatology differ in their thinking in a manner which determines differences in performance on subtests of the Wechsler-Bellevue. Nevertheless, the foregoing analyses imply that there is some association between the degree to which a patient's performance on the Wechsler test is impaired and the degree to which his general symptomatology represents a disturbance of behavior in general and of thought processes in particular. The pattern of correlations revealed by Table 2 suggests that, in general, patients with symptoms of conversion hysteria[3] and patients with a diagnosis of psychopathic personality[3] are the ones who tend to do best on the parts of the intelligence test, whereas the patients who are greatly excited[3] or show symptoms which we might consider hebephrenic[3] or deteriorated tend not to do well on the parts of the intelligence test. An aspect of test performance which may be characteristic of one type of patient is suggested by the negative correlation between depressed symptoms and performance on the arithmetic subtest of the Wechsler-Bellevue. This negative correlation is probably an expression of the delayed verbal responses so commonly observed among profoundly depressed patients.

The correlations are such as to suggest the possibility that the subtests of the Wechsler-Bellevue may be used to predict the hospitalized patients who show the greatest degree of conversion symptoms. Although the present sample is really too small to justify the com-

[3]The observed differences are not necessarily a result of the detracting affect of mental illness; the differences may express selective premorbid differences in mental ability; this possibility is contrary to the generally observed intellectual characteristics of hysterical and psychopathic groups, however.

putation of a multiple correlation coefficient, it is found that the correlation coefficient between scores on the conversion cluster and scores on the information, digit span, vocabulary, and picture completion subtests is .52.

CONCLUSION

From the data provided by this sample, it appears that the possibilities of using Wechsler-Bellevue subtest scores in differential diagnosis or in the prediction of the type of symptoms the patient will manifest are limited and may be so restricted as to be of no practical value. There is certainly little about the present data which would encourage the hope that these subtests could be used profitably in the descriptive diagnosis of an individual. It is possible that for certain types of group testing procedures or experiments, the degree of relationship revealed by these data would justify some attempt to use the subtest in predicting the symptom patterns revealed by the patients. In the clinic, however, the Wechsler-Bellevue is best regarded as an excellent test for intelligence and not as a major mechanical aid for descriptive diagnosis of functional disorders.

The results of the present analysis are of considerable incidental interest, however. They suggest that for research purposes the quantified multiple diagnostic procedure based on the use of rating scales may be used more profitably than psychiatrists' diagnoses. When the quantified multiple diagnostic procedure was designed, it was assumed that it would have numerous advantages over the use of psychiatrists' diagnoses. The difference between Analysis 1 and Analysis 2 in terms of the number and the strength of the relationships revealed lends support to this assumption. The research use of the quantified multiple diagnostic procedure not only has rational advantages over psychiatrists' diagnoses, but in view of the present results it has empirical advantages.

Received November 22, 1950.

REFERENCES

1. GARFIELD, S. L. A preliminary appraisal of Wechsler-Bellevue scatter patterns in schizophrenia. *J. consult. Psychol.*, 1948, **12**, 32–36.
2. LEVINE, L. S. The utility of Wechsler's patterns in the diagnosis of schizophrenia. *J. consult. Psychol.*, 1949, **13**, 28–31.
3. RAPAPORT, D. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1945. Vol. I.
4. SCHLOSSER, J. R., AND KANTOR, R. E. A comparison of Wechsler's deterioration ratio in psychoneurosis and schizophrenia. *J. consult. Psychol.*, 1949, **13**, 108–110.
5. SHOBEN, E. J., JR. The Wechsler-Bellevue in the detection of anxiety: A test of the Rashkis-Welsh hypothesis. *J. consult. Psychol.*, 1950, **14**, 40–45.
6. WITTENBORN, J. R. An evaluation of the use of Bellevue-Wechsler subtest scores as an aid in psychiatric diagnosis. *J. consult. Psychol.*, 1949, **13**, 433–439.
7. WITTENBORN, J. R. A new procedure for evaluating mental hospital patients. *J. consult. Psychol.*, 1950, **14**, 500–501.
8. WITTENBORN, J. R. An evaluation of the use of difference scores in prediction. *J. clin. Psychol.*, 1951, **7**, 108–111.

THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN AND THE REVISED STANFORD-BINET

ARTHUR WEIDER, PAUL A. NOLLER, AND THEODORE A. SCHRAMM
UNIVERSITY OF LOUISVILLE SCHOOL OF MEDICINE

CLINICAL psychodiagnostic practice is usually considerably enhanced when a new psychological device is standardized and becomes available to the clinician. The new scale is usually studied in relation to older tests in an attempt to evaluate it. Furthermore, in clinical practice it often becomes imperative not only to know to what degree any two instruments are correlated, if at all, but how to interpret divergent IQ's when two different scales have been used.

With the advent of the Wechsler Intelligence Scale for Children (WISC) for the age groups formerly served primarily by the Stanford-Binet, it was felt that such a correlational study should be made so that the clinician availing himself of either scale might readily compare scores, and in cases of discrepancy, seek additional data in order to arrive at the most reliable level of intelligence.

Similar studies are found in the literature, wherein the Stanford-Binet, Form L, 1937 Revision, is compared to the Wechsler-Bellevue Intelligence Scale for Adults [1, 3] and for Adolescents [2, 6]. In some of these studies, provisional scales of equivalent scores were computed for Wechsler-Bellevue IQ's from a knowledge of Stanford-Binet IQ's. The purpose of the present study is primarily designed to present similar data that should also prove useful in the psychological clinic.

METHOD

A group of 106 children comprising an age range from 5 years, 0 months, to 11 years, 11 months, was given both the WISC and Form L of the Revised Stanford-Binet. All of the subjects were white and lived in the city of Louisville. In order to determine the relationship of IQ scores of these two tests and the influence of age, the entire group of children was divided into two groups: Group A, consisting of 44 children who were 5 years, 0 months, to 7 years, 11 months, inclusive; and Group B, consisting of 62 children, 8 years, 0 months, to 11 years, 11 months, inclusive. In Group A, 23 of the subjects were boys and 21 were girls. In Group B, 38 of the subjects were boys and 24, girls. Most of the tests were given in two sessions, usually on the morning and the afternoon of the same day; however, many were administered in one sitting.

TABLE 1

CORRELATIONS OF SUBSCALES AND FULL SCALE OF THE WISC WITH FORM L OF THE REVISED STANFORD-BINET

Group	Binet vs. WISC Full Scale	Binet vs. WISC Verbal Scale	Binet vs. WISC Perf. Scale
A (5–0 to 7–11)	.90 ± .03	.82 ± .05	.79 ± .05
B (8–0 to 11–11)	.90 ± .02	.92 ± .02	.78 ± .05
All Subjects	.89 ± .02	.89 ± .02	.77 ± .04

RESULTS

The coefficients of correlation between the several WISC scales and the Stanford-Binet are given in Table 1. For the most part, these correlations are the same for Group A, the younger subjects, and for Group B, the older subjects, except when the verbal scale of the WISC is compared to the Binet for the younger children. That the performance scale does not correlate as well with the Binet as does the verbal scale is to be expected in the light of the lesser proportion of such material in the Binet. This same trend, of the performance correlation being lower than the verbal, has been found with the adult Wechsler scale and the Binet [1, 2].

TABLE 2
MEANS AND STANDARD DEVIATIONS OF BINET AND WISC IQ SCORES

Scale	All Subjects Mean	SD	Range	Group A (5–0 to 7–11) Mean	SD	Range	Group B (8–0 to 11–11) Mean	SD	Range
Binet	93.1	19.56	87(51–138)	100.0	17.75	72(62–134)	88.2	19.43	87(51–138)
WISC Full	90.0	18.90	83(46–129)	92.7	18.60	68(52–120)	88.1	19.03	83(46–129)
WISC Verbal	91.9	17.07	73(56–129)	94.8	17.98	73(56–129)	89.9	16.23	63(65–128)
WISC Perf.	89.8	18.14	85(44–129)	91.7	17.76	66(55–121)	88.4	18.19	85(44–129)

In Table 2 are listed mean IQ scores, standard deviations, and ranges of scores of the total population and of the subgroups. It will be noted that the mean scores for the scales, with the exception of the performance mean scores, are relatively similar for each population group. The range of IQ scores of the Binet does not vary too much from that of the other scales, either for total population or for separate age groups.

TABLE 3
COMPARATIVE TEST SCORES OF SUBJECTS MAKING MAXIMAL AND MINIMAL BINET AND WISC SCORES

Highest Binet IQ's Subject	Binet	WISC Full Scale	Lowest Binet IQ's Subject	Binet	WISC Full Scale
A	138	114	G	51	46
B	135	129	H	62	64
C	134	120	I	66	69
D	132	113	J	71	64
E	131	120	K	71	75
F	129	101	L	72	52

Highest WISC IQ's Subject	WISC Full Scale	Binet	Lowest WISC IQ's Subject	WISC Full Scale	Binet
C	120	134	G	46	51
B	129	135	L	52	72
M	125	128	P	62	75
N	123	122	Q	64	87
O	123	126	H	64	62
P	120	116	J	64	71

In Table 3, the Binet and WISC scores are compared for subjects who made the highest and lowest scorre on each scale. It is seen that there is no pattern of consistency for either high or low scores. It would appear that the subjects who make the highest Binet scores have higher Binet than WISC IQ's. Those whose WISC scores are highest show no such pattern. The same observations seem to hold for the comparison of the low scores: those whose Binets were lowest show no consistency, while those whose WISC scores were lowest, in four of the six cases cited, had scores below those of the corresponding Binet.

TABLE 4
EQUIVALENT STANFORD-BINET AND WISC FULL SCALE IQ SCORES*
P.E. = 5.8 IQ POINTS
(EST. BELLEVUE IQ)

Binet IQ	WISC IQ	Binet IQ	WISC IQ	Binet IQ	WISC IQ
40	45	74	74	108	103
42	47	76	76	110	105
44	48	78	77	112	106
46	50	80	79	114	108
48	52	82	81	116	110
50	53	84	82	118	111
52	55	86	84	120	113
54	57	88	86	122	115
56	59	90	88	124	116
58	60	92	89	126	117
60	62	94	91	128	120
62	64	96	93	130	122
64	65	98	94	132	123
66	67	100	96	134	125
68	69	102	98	136	127
70	71	104	99	138	128
72	72	106	101	140	130

*Based on the regression equation $y = 0.85x + 11$, where "y" is the WISC IQ score and "x" is the Binet IQ score.

These differences are reflected in the regression equation ($y = 0.85x + 11$) for predicting the WISC IQ from the Binet IQ. This regression equation is for the full scale IQ for both age groups combined. Table 4 presents the provisional "equivalent scores." The coefficients of variation were calculated as follows:

a) Total population—for WISC, 20.23 per cent; for Binet, 18.48 per cent
b) For Group A (5–0 to 7–11 years)—for WISC, 12.42 per cent; for Binet, 24.64 per cent
c) For Group B (8–0 to 11–11 years)—for WISC, 15.25 per cent; for Binet, 18.65 per cent

Of these, only the figures for the younger group suggest any significance, indicating that Binet scores are subject to more dispersion for this age group than are the WISC scores.

DISCUSSION

It should be borne in mind that the data presented here have significance in so far as the performance of normal children is compared on these two clinical instruments. In all probability, different results would be found in a similar study on children with emotional problems. However, it is always a decided advantage to the clinician to have at his disposal one or more similar instruments that tap, perhaps, different facets of intellective functioning. To be sure, there are advantages and disadvantages to either scale under discussion; however, these can be minimized or maximized depending upon specific circumstances. It is not within the scope of this paper to investigate which tool measures a child's intelligence with greater validity, but rather to show the relationships between their scores. The great success enjoyed by the Wechsler-Bellevue Intelligence Scale for Adults no doubt recommends the careful application of the WISC. To be sure, some of the advantages of the Bellevue for adult populations pertain also to the measurement of the intelligence of children. Thus the separation of the Bellevue Scale into subtests enables the continuous flow of test objects to be placed before the subject, perhaps thereby insuring a greater degree of interest to children. In addition, the psychodiagnostic qualities offered by the WISC will recommend its use in many clinical situations.

Wechsler [5] and Seashore, Wesman, and Doppelt [4] make the point that the children's scale is quite distinct from the adult Wechsler scales since it has been independently standardized. There is some similarity of test material between the children's and the adult scales, but according to Wechsler, the same test items have differing psychological value and meaning for adults and children. In the WISC, a given child's test performance is compared with the scores of other children of his own age group rather than with scores drawn from a composite age group.

The tremendous literature on the Stanford-Binet, the amount of experience with the scale, and the large body of information obtained with it, naturally recommend its continued application in clinical practice. It is because of this continued interest that the writers have sought to prepare Table 4, which should be helpful in quickly ascertaining the statistical relationship that exists between these two well-standardized clinical instruments. The writers look forward to similar studies in the future on many varieties of clinical and nonclinical subjects, studies which should reveal significant information for clinical psychodiagnostic practice.

SUMMARY

The Wechsler Intelligence Scale for Children and Form L of the Revised Stanford-Binet Intelligence Scale were administered to a group of 106 white Louisville children, ranging in age between 5 years, 0 months, and 11 years, 11 months. The IQ scores obtained were correlated for a younger group (5–0 to 7–11 years) and for an older group (8–0 to 11–11 years) as well as for all the subjects. Correlations were obtained between the verbal scale, performance scale, and full scale of the WISC and the Stanford-Binet. Studies of the statistical results yielded the following:

1. The coefficient of correlation for all subjects between the full scale WISC IQ and the Binet IQ is .89 ± .02; for the verbal scale and Binet, .89 ± .02, and for the performance scale and the Binet, .77 ± .04.

2. Binet IQ's tend to be higher than the WISC IQ's for the same children. This is borne out by the table of equivalent scores derived from the regression equation.

3. A provisional table of "equivalent scores" for the prediction of the WISC IQ from Binet IQ scores, based on the population studied, is included.

Received November 10, 1950.

REFERENCES

1. BENTON, A. L., WEIDER, A., AND BLAUVELT, J. Performances of adult patients on the Bellevue intelligence scales and the Revised Stanford-Binet. *Psychiat. Quart.*, 1941, **15**, 802–806.

2. HALPERN, F. A comparison of the Revised Stanford L and the Bellevue Adult Intelligence Test as clinical instruments. *Psychiat. Quart. Suppl.*, 1942, **16**, 206–211.
3. MITCHELL, M. B. Performance of mental hospital patients on the Wechsler-Bellevue and the Revised Stanford-Binet, Form L. *J. educ. Psychol.*, 1942, **33**, 538–544.
4. SEASHORE, H., WESMAN, A., AND DOPPELT, J. The standardization of the Wechsler Intelligence Scale for Children. *J. consult. Psychol.*, 1950, **14**, 99–110.
5. WECHSLER, D. *Wechsler Intelligence Scale for Children.* New York: Psychological Corp., 1949.
6. WEIDER, A., LEVI, J., AND RISCH, F. Performances of problem children on the Wechsler-Bellevue Intelligence Scales and the Revised Stanford-Binet. *Psychiat. Quart.*, 1943, **17**, 695–701.

THE DEVELOPMENT OF A SZONDI INSTABILITY SCORE

HENRY P. DAVID
TEACHERS COLLEGE, COLUMBIA UNIVERSITY
AND
WILLIAM RABINOWITZ
DIVISION OF TEACHER EDUCATION, COLLEGE OF THE CITY OF NEW YORK

IT is generally assumed by those working with the Szondi Test [4] that it reflects the personality as an ever-changing process. According to Deri [3], the first step in the interpretation of a Szondi profile consists of a general inspection of the entire series of tests administered to a given subject to "note a constancy or change in the modes of choice reactions." She maintains that "the various degrees and types of changes are the first indicators which help us discriminate between the main categories of normal or abnormal behavior."

Although there are several ways in which the stability of Szondi responses could be quantified, none has appeared in the literature thus far. It is apparent that scores might be developed based upon changes in factors, or upon changes in reaction to the individual Szondi pictures. The second procedure seems to offer several advantages. Each of the pictures can be scored separately, and the results may be combined to yield factor scores as well as a total score. This method has been utilized in the development of a Szondi Instability Score, hereafter referred to as SIS.

METHOD AND PROCEDURE

The SIS is based upon a subject's reactions to each of the 48 Szondi photographs over a given series of administrations. Score values are assigned to changes in reaction from one administration to the next according to the following scale:

0 No change in reaction
1 From "liked" to "ignored," or from "disliked" to "ignored," or vice versa
2 Reversal from "liked" to "disliked," or vice versa

The scoring procedure requires a separate unit score for each Szondi picture. These are subsequently added together to provide the SIS. The SIS is therefore, in part, a function of the number of administrations being scored.

The Szondi Test was administered, individually, six times, to each of 20 hospitalized paranoid schizophrenic women and to each of 20 student nurses.[1] The latter represented a fairly homogeneous group, none of whom had ever been under psychiatric care. Administration procedure and scoring followed the directions outlined in Deri's volume [2]. In no instance was the time interval between administrations less than 24 hours. All the patients and all the nurses were unmarried. Changes in reaction to each of the 48 photographs were recorded and scored according to the SIS value scale. The data have been summarized in Table 1.

RESULTS AND DISCUSSION

Inspection of Table 1 indicates that differences in the factor scores *within* each of the two groups are quite small. Apparently the separate factors do not differ markedly in their respective contributions to the total measured instability of the individual groups. Examination of the factor scores *between* the two

[1]The paranoid schizophrenic group was tested at Longview State Hospital, Cincinnati, Ohio, with the permission of E. A. Baber, M.D. Superintendent, and with the aid of Doris Twitchell-Allen, Ph.D., Chief Psychologist. The student nurses were seen at Good Samaritan Hospital, Cincinnati, Ohio, with the cooperation of Howard D. McIntyre, M.D.

TABLE 1
SIS SCORE FOR 20 PARANOID SCHIZOPHRENICS AND 20 STUDENT NURSES

	Factor	Instability Scores																				Total
Student Nurses	*h*	0	3	4	4	7	3	7	8	8	5	5	6	13	9	6	3	7	5	16	12	131
	s	1	0	2	6	4	2	7	3	1	6	7	9	5	9	8	9	13	15	12	12	131
	e	1	0	4	7	4	8	3	8	4	4	10	3	14	10	14	5	13	12	15	15	151
	hy	4	3	4	0	3	5	6	5	4	5	8	12	4	8	7	5	4	13	8	6	114
	k	0	0	7	8	8	4	4	6	8	10	1	8	5	8	13	13	10	10	5	15	143
	p	2	0	5	6	5	11	10	4	6	10	10	10	5	8	9	11	10	4	7	10	143
	d	1	3	4	1	3	5	7	8	9	7	10	8	8	10	9	12	8	8	12	7	140
	m	3	3	2	4	9	6	5	8	10	5	5	4	13	8	6	17	14	12	12	11	157
Total (SIS)		12	12	32	36	43	44	49	50	50	52	56	60	67	70	72	75	79	79	84	88	1,110
Paranoid Schizophrenics	*h*	8	8	3	7	9	12	8	19	14	7	16	4	13	14	13	16	16	22	22	21	252
	s	5	5	3	10	15	9	11	7	12	18	8	15	18	14	11	17	26	12	18	19	253
	e	5	2	6	7	9	11	11	15	9	16	12	20	7	19	25	18	19	18	25	26	280
	hy	4	7	3	12	7	15	10	4	11	12	14	15	18	16	16	17	23	13	19	29	265
	k	2	1	10	13	14	6	12	15	13	13	11	11	11	16	15	17	18	19	13	18	248
	p	8	4	2	14	9	12	11	11	15	18	16	14	21	14	19	18	19	25	15	23	288
	d	1	5	7	9	11	8	12	15	16	11	12	18	17	12	14	11	13	22	19	18	251
	m	0	4	4	4	4	13	15	8	10	9	21	17	13	15	13	24	14	25	27	22	262
Total (SIS)		33	36	38	76	78	86	90	94	100	104	110	114	118	120	126	138	148	156	158	176	2,099

groups, however, reveals large differences. The SIS's of the individual subjects show similarly great differences. The paranoid schizophrenic women obtained a mean SIS of 105 and a variance of 1,646.3, whereas the student nurses obtained a mean score of 55.5 and a variancce of 479.4. Statistical comparison of the means and variances yield a t of 3.03 and an F of 3.43. Both of these values are significant at the .01 level.

While a cut-off SIS cannot be validated solely on the basis of the data gathered here, it may be of interest to hypothesize such a score. Examination of the SIS's obtained indicates that 90 might be selected as a tentative cut-off score. All of the student nurses obtained an SIS lower than 90, while 70 per cent of the paranoid schizophrenic women obtained an SIS above this score. It is, of course, recognized that before any hypothesized cut-off point can be accepted it must be cross-validated on an independent sample of the population.

The test profiles of the paranoid schizophrenic women had been selected at random from a larger experimental group [1]. It is of interest that the highest scores were obtained by the patients residing in the most disturbed hospital ward. Those women who were relatively less disturbed generally obtained lower scores.

At this stage of development the SIS should be employed with caution. Other groups should be examined before its clinical value can be fully assessed and evaluated.

SUMMARY

The Szondi Test was administered to 20 hospitalized paranoid schizophrenic women and to 20 student nurses. A Szondi Instability Score, SIS, based upon changes in choice reactions to each of the 48 photographs was developed. This technique differs from the usual method of Szondi profile interpretation and analysis of changes in factorial constellations. The paranoid schizophrenics obtained a higher mean and greater variance on the SIS than did the student nurses. The clinical meaning of the SIS was considered.

Received November 1, 1950.

References

1. David, H. P. An evaluation of repeated administrations of the Szondi Test. Unpublished Master's thesis, Univ. of Cincinnati, 1949.
2. Deri, Susan. *Introduction to the Szondi Test.* New York: Grune & Stratton, 1949.
3. Deri, Susan. The Szondi Test. *Amer. J. Orthopsychiat.*, 1949, **19**, 447–454.
4. Szondi, L. *The Szondi Test.* New York: Grune & Stratton, 1948.
5. Szondi, L. *Experimentelle Triebdiagnostik.* New New York: Grune & Stratton, 1947.

THE INFLUENCE OF SET FOR AREA ON THE RORSCHACH TEST RESULTS[1]

LEONARD S. ABRAMSON
VETERANS ADMINISTRATION HOSPITAL, KNOXVILLE, IOWA[2]

THE Rorschach test, as an unstructured method of personality study, evaluates personality by the manner in which the subject perceives and structures standardized ink blots. The test instructions are purposely devoid of suggestions as to what constitutes an adequate performance. The subject, left to his own devices in this novel situation, must assume his own standards of adequacy. He derives these primarily from his personality characteristics. Standards may also be learned from external sources or influences, suggestions, hearsay, variations in the testing conditions,, etc. Used in employment selection procedures, for example, how impervious would the test be to the influence of such uncontrolled factors?

The authors of the two leading manuals on the Rorschach technique, recognizing the difficulty of presenting formal test instructions for this relatively unstructured test, do not give express instructions for its administration. Beck [1] states that no formula is to be used verbatim and that paraphrasing may be "very liberal" as long as "no departure in essence" is made from the given formula. Klopfer [9] is less explicit. He acknowledges that there is quite a difference of opinion as to how fully the subject should be told what to do. Rather than stating explicit instructions as does Beck, Klopfer presents instructions compounded of three formulations with the suggestion that if the subject tends too greatly in any one direction, he should be reminded of the other two. Formal instructions which will orient the subject's approach have not as yet been specified.

As noted, formal instructions are not the only ones present in any test situation. McGeoch [10] cites three kinds which are inherent in all tests; formal instructions, incidental or occasional instructions, and self-instructions. Self-instructions might include a variety of test attitudes such as those discussed by Schachtel [12], "subjective definitions," based on attitudes toward authority, competition and resistance. These attitudes do not hamper the test results but, in their discovery through the test, contribute to an understanding of the subject.

Several experiments have dealt with the effects of differential test instructions on the Rorschach. Fosberg [3, 4] found that neither explanation to the subjects of all of the test factors, nor asking them to produce "good" and "bad" records, altered the test results. Single variables, however, were not controlled systematically. Coffin [2, pp. 47–65] set separate groups of subjects to perceive the ink blots as a whole and in terms of details. The subjects responded in the desired directions.

What would the effects of self-instructions, attitude, or "set" be on the protocol as a whole, on the personality pattern that a subjects presents? To be meaningful, the instructions should be ones which the subjects might themselves bring to the Rorschach test situation as self-instructions prior to any given them by the examiner. Sets about area (where and how many) and content would be ones which the subjects might most likely know. In

[1]This study was done in partial fulfillment of requirements for the Ph.D. degree. The writer is indebted to Dr. C. A. Whitmer, of the University of Pittsburgh, under whose direction the present study was completed, and to the University of Nebraska faculty members, who aided in obtaining the samples of subjects.

[2]Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the author are the results of his own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

order to study experimentally the effects of set by instructions, one of these aspects was selected, the area or location of the test responses. A number of questions were raised:

1. What effects would sets given by formal instructions to perceive the ink blots as a whole in one group, and in terms of their details or parts in another, have on the number and percentage of *W*, *D*, and *Dd* responses?

2. What effects would such sets have on the other scoring categories, the determinants and content?

3. What effects would these sets have on the adjustment scores obtained on the Rorschach test records, as measured by the Munroe inspection technique.

4. How consistent are the changes in the directions of the responses of the various Rorschach test components in the experimentally "set" groups, as compared with the test-retest reliability coefficients as measured on a control group?

SUBJECTS AND PROCEDURE

Volunteer subjects were solicited from classes in psychology and sociology at the University of Nebraska in the fall semester, 1948. One hundred and forty subjects, divided into three groups, were given a standard Rorschach test.[3]

After a three-week interval, the 122 subjects who returned for the retest were again tested in three groups, divided as follows: control group, 40 subjects; experimental group *W*, 45 subjects; and experimental group *D*, 37 subjects. The groups were equated on the relative amounts of *W*, *D*, and *Dd* responses given on the first test, and on sex, age, ACE Psychological Examination for College Freshmen L and Q scale stanine scores, years in college, and number of courses taken in psychology. The subjects were relatively naive as to psychological testing, almost three-quarters of the sample having had no more than an introductory psychology course. The distributions of the groups on the equating factors are given in Table 1. The mean, *W*, *D*, and *Dd* scores and significance of the differences between the groups are given in Table 2.

TABLE 1

EQUATION OF THE THREE GROUPS AS TO SEX, AGE, ACE SCORES, YEAR IN COLLEGE, AND NUMBER OF COURSES TAKEN IN PSYCHOLOGY

	Control	W Group	D Group
Number of Subjects	40	45	37
Males	21	24	21
Age in Years			
Mean	22.8	21.8	21.6
SD	7.1	4.3	2.8
ACE L Scale Stanine Score			
Mean	5.50	5.91	5.41
SD	1.61	1.24	1.81
ACE Q Scale Stanine Score			
Mean	4.77	5.67	4.64
SD	1.81	1.49	1.91
% of Sophomores and Juniors	80.0	60.0	70.2
% 0–1 Psychology Course	77.5	71.1	62.2

The tests were given in four sessions (two groups comprising experimental group *W*). Sender's instructions and "minimum prodding" inquiry were used [13]. Following the reading of the instructions the blots were exposed on a screen for two minutes each. The subjects wrote their responses in mimeographed booklets which were especially prepared. During the inquiry period, the blots were again shown for three minutes each while the subjects circled the areas of their responses on small blots and wrote descriptions of what they had seen.

The retests, given after a three-week interval, were administered in the same manner. The two experimental groups, however, were given additional "information" about the test. Each of the latter groups was told that successful business and professional people tended to perceive the blots in certain ways, which approaches to the test correlated positively with high intelligence. Experimental group *W* was set to perceive wholes, experimental group *D* to perceive details.

[3]The group Rorschach was utilized in order to test the large number of subjects under carefully controlled conditions as to knowledge about the test and "set." Studies by Hertz [7], Harrower-Erickson and Steiner [5], Sender [13], and others have shown results of group and individual methods to be comparable.

The scoring combined procedures and symbols from Beck, Klopfer, Hertz [6], and Munroe [11]. Beck's notations of area and Z score were used. The determinants of the perceptions were scored according to Klopfer because of the finer analysis which his system offers. Accuracy was scored on the basis of Hertz's tables, inasmuch as they were based on a statistical study of a defined normative

TABLE 2

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE MEANS OF THE SCORING CATEGORIES ON THE STANDARD TEST USED TO EQUATE THE THREE GROUPS FOR AREA

	Control	W Group	D Group	C–W Diff	CR	C–D Diff	CR*
A. Area							
N	30.80	30.44	30.41	−0.36	0.19	−0.39	0.20
W	8.08	8.38	8.11	+0.30	0.33	+0.03	0.04
W%	28.02	30.91	28.92	+2.89	0.94	+0.90	0.25
D	19.78	18.67	18.97	−1.11	0.68	−0.81	0.49
D%	63.65	60.34	61.05	−3.31	1.23	−2.60	0.90
Dd	2.95	3.40	3.33	+0.45	0.62	+0.38	0.48
Dd%	8.33	10.64	10.42	+2.31	1.28	+2.09	1.09
D+Dd	22.73	22.07	22.30	−0.66	0.32	−0.43	0.20
Wtd.S	1.33	1.11	1.51	−0.22	0.67	+0.51	0.43
Z	51.06	50.69	50.05	−0.37	0.09	−1.01	0.25
VIII–X%	38.00	36.71	36.19	−1.29	0.98	+1.81	1.16
B. Determinants							
M	5.53	5.49	5.70	−0.04	0.07	+0.17	0.25
W>M	2.60	2.89	2.38	+0.29	0.29	−0.22	0.22
FM	5.03	4.33	4.18	−0.70	1.28	−0.85	1.60
m	1.15	0.91	1.32	−0.24	1.09	+0.17	0.65
FC	2.53	3.41	3.03	+0.88	2.10	+0.50	1.16
CF	1.73	1.69	1.64	−0.04	0.13	−0.09	0.26
Sum C	3.43	4.14	3.72	+0.71	1.57	+0.29	0.68
M>C	2.05	1.39	2.04	−0.66	1.00	−0.01	0.01
Fc	2.15	2.24	2.23	+0.09	0.26	+0.08	0.24
cF	0.54	0.51	0.35	−0.03	0.17	−0.19	1.19
FC′	0.59	0.71	0.46	+0.12	0.63	−0.13	0.76
C′F	0.19	0.17	0.20	−0.02	0.08	+0.01	0.04
FM+m> Fc+c+C′	2.48	1.44	2.03	−1.04	1.25	−0.45	0.60
FK	1.36	1.02	1.30	−0.34	1.23	−0.06	0.20
% of "blends"	15.91	15.30	15.35	−0.61	0.29	−0.56	0.29
F% (Klopfer)	35.40	36.64	37.76	+1.24	0.46	+2.36	0.85
F>Fc+FK	7.05	7.24	7.41	+0.19	0.16	+0.36	0.30
Beck F+%	78.47	82.98	78.85	+4.51	1.88	+0.38	0.15
Total F+%	83.37	84.44	83.08	+1.07	0.69	−0.29	0.17
C. Content							
A%	44.58	42.07	41.59	−2.51	1.01	−2.99	1.35
H%	21.23	23.29	23.14	−2.06	0.94	+1.91	0.96
H+A%	65.78	65.11	64.73	−0.67	0.27	−1.05	0.41
A>H	6.98	5.87	5.51	−1.11	0.93	−1.47	1.41
Hd+Ad%	16.34	18.40	16.97	+2.06	0.91	+0.63	0.28
A+H> Ad+Hd	9.00	8.58	8.97	−0.42	0.31	−0.03	0.02
An	1.35	1.50	1.00	+0.15	0.52	−0.35	1.30
Oj	1.00	1.00	1.16	0.00	0.00	+0.16	0.60
Beck P	7.65	7.33	6.95	−0.32	0.81	−0.70	1.73
Klopfer P	5.86	5.28	5.31	−0.58	1.93	−0.55	1.83

*CR if 1.96 and beyond, significant at the 5 per cent level of confidence.

CR if 2.57 and beyond, significant at the 1 per cent level of confidence.

TABLE 3

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE MEANS OF THE SCORING CATEGORIES ON THE RETEST UNDER STANDARD AND EXPERIMENTALLY "SET" CONDITIONS

	Control	W Group	D Group	C–W Diff	CR	C–D Diff	CR*
A. Area and Related Components							
N	35.70	29.58	45.27	−6.12	3.09	+9.57	3.94
W	6.10	12.73	4.70	+6.63	6.70	−1.40	2.00
W%	18.66	44.67	11.09	+26.01	8.79	−7.57	3.54
D	24.60	15.13	31.70	−9.47	6.31	+7.10	4.18
D%	68.60	50.04	70.46	−18.56	7.40	+1.86	1.12
Dd	5.00	1.71	8.86	−3.29	4.38	+3.86	3.51
Dd%	12.74	5.28	18.54	−7.46	4.87	+5.80	2.94
D+Dd	29.60	16.84	40.57	−12.76	6.25	+10.97	4.35
Wtd.S	1.74	1.10	2.96	−0.64	1.56	+1.22	2.00
Z	47.13	60.82	46.38	+13.69	4.14	−0.75	0.20
VIII–X%	37.30	37.24	36.03	−0.06	0.05	−1.27	1.06
B. Determinants							
M	6.23	5.37	6.74	−0.86	1.55	−0.51	0.62
W>M	0.10	7.37	−2.04	−7.47	6.79	−1.94	1.83
FM	5.25	3.76	5.96	−1.49	2.62	−0.71	1.05
m	1.24	1.17	2.12	−0.07	0.30	−0.88	2.98
FC	3.54	4.27	4.27	−0.73	1.46	−0.73	1.20
CF	1.45	1.89	1.77	−0.44	1.49	−0.32	0.95
Sum C	3.63	4.50	4.46	−0.87	2.03	−0.83	1.62
M>C	2.64	0.82	2.88	−1.82	2.76	−0.26	0.81
Fc	2.68	2.77	3.74	−0.09	0.24	−1.06	2.52
cF	0.44	0.59	0.69	−0.15	0.88	−0.25	1.27
FC′	0.98	0.91	1.11	−0.07	0.30	−0.13	0.45
C′F	0.19	0.30	0.18	−0.11	0.39	−0.01	0.04
FM+m> Fc+c+C′	1.65	0.14	1.95	−1.51	1.67	+0.30	0.33
FK	1.26	1.08	1.77	−0.18	0.60	+0.51	1.34
% of "blends"	15.77	18.66	14.97	+2.89	1.35	−0.80	0.35
F% (Klopfer)	37.48	30.51	41.27	−6.97	2.48	+3.79	1.32
F>Fc+FK	9.73	4.53	12.24	−5.20	3.47	+2.51	1.45
Beck F+%	83.41	81.02	83.19	−2.39	1.03	−0.22	0.11
Total F+%	85.60	84.51	84.62	−1.09	0.79	−0.98	0.71
C. Content							
A%	44.33	38.71	38.92	−5.62	2.35	+5.41	2.56
H%	22.10	22.82	21.86	+0.72	0.35	−0.24	0.12
H+A%	66.43	61.53	60.78	−4.90	2.52	−5.65	3.14
A>H	7.35	4.98	7.84	−2.37	1.71	+0.49	0.33
Hd+Ad%	18.39	15.27	21.57	−3.12	1.63	+3.18	1.47
A+H> Ad+Hd	9.85	9.07	7.59	−0.78	0.57	−2.26	1.45
An	0.78	0.80	0.92	+0.02	0.11	+0.14	0.61
Oj	1.53	1.11	2.05	−0.42	1.64	+0.52	1.72
Beck P	8.08	7.22	7.41	−0.86	2.13	−0.67	1.57
Klopfer P	6.23	4.96	5.53	−1.27	3.85	−0.70	2.26

*CR if 1.96 and beyond, significant at the 5 per cent level of confidence.

CR if 2.57 and beyond, significant at the 1 per cent level of confidence.

sample. Use of the Munroe inspection technique entailed additional quantitative and qualitative aspects of scoring. A total of 84 discrete scoring factors and ratios were analyzed, including 27 components of the Munroe adjustment score. Some of these 84 factors are presented in Tables 2 and 3. Table 4 presents a comparison of the over-all Munroe adjustment scores.

RESULTS

A. *Area.* Comparisons of the three groups on the first test are presented in Table 2, on the retest under control and experimental conditions in Table 3.

Both sets effected highly significant shifts of the various area designations of responses in the directions which were suggested. Differences between the groups were insignificant on the first test; many were significant on the retest.

Experimental group *D* gave evidence of a quantitative drive. Their records increased in the average number of responses from 30.41 to 45.27 responses. Experimental group *W* decreased slightly in number of responses and their responses indicated greater concern with quality of response. The findings tend to confirm the results of Coffin's study on suggestibility.

B. *Other Scoring Categories.*

1. Experimental Group *W*. Statistically significant changes, as compared with the control group, were found on a number of the non-*F* determinants including blends of these determinants. The *F* percentage decreased significantly. Movement responses, except for inanimate movement, were relatively unaffected. Being more sensitive to color and using proportionately more despite the decrease in total number of responses, this group showed a great susceptibility to color shock as evaluated in the Munroe technique. In content, stereotypy decreased. A number of content categories showed significant increases.

2. Experimental Group *D*. As in the other experimental group, *M* responses were little affected. An increase of *FM* and especially *m* responses resulted in a more maladjustive relationship between the various movement responses. Color shock, contrary to experimental group *W*, decreased. Shading shock, evaluated for the Munroe technique, however, increased significantly relative to the other two groups. As expected, form percentage and form accuracy increased. Stereotypy decreased in this experimental group also as the subjects sought new responses. *Hd* and *Ad* responses increased. The results were indicative of increased maladaptiveness.

TABLE 4

SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE MEAN MUNROE INSPECTION TECHNIQUE ADJUSTMENT SCORES ON THE FIRST AND SECOND TESTS

	First Test		Second Test	
	Mean	SD	Mean	SD
Control	7.75	2.20	7.45	3.40
W Group	7.87	3.44	6.76	3.52
D Group	7.81	3.27	9.97	5.36
C–W Diff	+0.12		—0.69	
CR	0.19		0.91	
C–D Diff	+0.06		+2.52	
CR	0.09		2.42*	
W–D Diff	—0.06		+3.21	
CR	0.08		3.09*	

*Significant at the 1 per cent level of confidence

C. *Munroe Inspection Technique.* The over-all results on the Munroe technique, given by the adjustment score, corroborate the individual components cited above. Table 4 indicates the salutary effect of the set to perceive wholes and the maladjustive effect of the set to perceive details.

D. *Test-Retest Correlations.* The median test-retest correlations (*r*'s) obtained on the control group were: area, .75; determinants, .62; content, .73. The experimental sets effected a decrease in median test-retest correlations as the subjects sought varying and apparently disparate methods of meeting the "set" conditions. The correlations were, respectively, .51, .52, and .55 on experimental group *W*, and .68, .62, and .54 on experimental group *D*.

Space does not permit the inclusion of all the correlations obtained. Hutt and others [8] have presented test-retest reliability coefficients on individual tests given to subjects ($N = 21$) at the University of Michigan. The correlations given by Hutt (rank order, ρ, and Pearsonian r) are of a comparable mag-

nitude with those obtained on the present sample of 40 in the control group: *W* (Hutt's cited first), ρ .62, *r* .78; *D* + *Dd*, ρ .62, *r* .69; *R*, ρ .83, *r* .64; *M*, ρ .66, *r* .84; *S*, ρ .66; *r* .84; *H%*, *r* .87, *r* .85; % color responses, *r* .66, Sum *C*, *r* .51; and non-*F%*, *r* .81, *F%*, *r* .62. Three components differed considerably, correlating high in the present sample, but low on that of Hutt: *A%*, ρ .37, *r* .78; *F+%*, *r*—.13, *r* .54; and *FC*, *r* .08, *r* .58. The Rorschach test proved to be of a fairly high degree of reliability in our study, except where an intervening experimental set was introduced.

DISCUSSION

The present study suggests that the Rorschach test results may be significantly altered by set or suggestion. The sets given by formal instructions were such as subjects might themselves assume and apply to the test by self-instructions.

The findings give further indication of the extent to which perception may be altered by changed motive-incentive conditions. The significant changes resulting from the conscious efforts to perceive the blots in the manners suggested also represents the effects of increased ego-involvement. The subjects sought to approximate standards which were defined. Whether abnormal subjects would be able to do so remains to be evaluated. We would assume in agreement with Hutt [8] that the ability to "shift" might distinguish between pathological individuals and those of the normal population. A Rorschach psychogram would therefore be less reliable in the case of a normal subject.

The more immediate implications relate to the use of the Rorschach test in practice. The importance of fixed and standardized instructions is indicated by the extent to which records may be altered by varied instructions. Testing of the limits is particularly indicated for the purpose of learning the subject's possible preconceptions. Certain of the test components, particularly *M* determinants and *H* content, are relatively resistant to altered set. Rorschach interpretation, at the psychogram level, should be based upon such components.

SUMMARY

This study investigated the effects on the Rorschach test results of set given by instructions to perceive the ink blots as a whole in one experimental group and in detail in another experimental group.

Three groups of college subjects, a total of 122 students, were equated as to age, sex, ACE scores, years in college, number of courses in psychology, and the area of their responses on a standard group Rorschach pretest. After a three-week interval, the three groups were retested. One group served as a control group for test-retest changes, and the other two groups served as the experimental *W* and *D* groups.

The following results were obtained:

1. All components of the areas of response changed significantly in the directions of the two experimental sets.
2. Many of the determinant and content components and ratios were also altered by the experimental sets for the area or location only of the responses.
3. On the Munroe inspection technique, the mean adjustment score of the control group was unaltered from test to retest, that of the *W*-set experimental group decreased, and that of the *D*-set group increased significantly toward a more "maladjusted" score.
4. The median test-retest reliability coefficients of the two experimental groups decreased relative to those of the control group.

The results of this study suggest the importance of standardized instructions on the Rorschach test and a careful testing of the limits to ascertain what the subject's preconceptions were prior to taking the test. As Hutt suggests, testing of a subject under different sets might be utilized to measure capacity to "shift" as a means of evaluating normal flexibility as against the fixity and rigidity found in pathological cases.

Received December 1, 1950.

REFERENCES

1. BECK, S. J. *Rorschach's test: I. Basic processes.* New York: Grune & Stratton, 1944.

2. COFFIN, T. E. Some conditions of suggestion and suggestibility. *Psychol. Monogr.*, 1941, 53, No. 4.
3. FOSBERG, I. A. An experimental study of the reliability of the Rorschach psychodiagnostic technique. *Rorschach Res. Exch.*, 1941, 5, 24–36.
4. FOSBERG, I. A. How do subjects fake results on the Rorschach tests? *Rorschach Res. Exch.*, 1943, 7, 119–121.
5. HARROWER-ERICKSON, MOLLY R. Modification of the Rorschach method for use as a group test. *Rorschach Res. Exch.*, 1941, 5, 130–144.
6. HERTZ, MARGUERITE. *Frequency tables to be used in scoring responses to the Rorschach ink-blot test.* (3rd Ed.) Cleveland: Western Reserve Univ. Press, 1946.
7. HERTZ, MARGUERITE. Modification of the Rorschach ink-blot test for large scale application. *Amer. J. Orthopsychiat.*, 1943, 13, 191–212.
8. HUTT, M. L., GIBBEY, R., MILTON, E. O., AND POTTHARST, K. The effect of varied experimental "sets" upon Rorschach test performance. *Rorschach Res. Exch.*, 1950, 14, 181–187.
9. KLOPFER, B., AND KELLEY, D. McG. *The Rorschach technique.* Yonkers, N.Y.: World Book Co., 1946.
10. MCGEOCH, J. A. *The psychology of human learning.* New York: Longmans, Green, 1942.
11. MUNROE, RUTH L. Prediction of the adjustment and academic performance of college students by a modification of the Rorschach method. *Appl. Psychol. Monogr.*, No. 7, Stanford, Calif.: Stanford Univ. Press, 1945.
12. SCHACHTEL, E. G. Subjective definitions of the Rorschach test situation and their effect on test performance. *Psychiatry*, 1945, 8, 419-448.
13. SENDER, SADIE. The influence of variations in the Rorschach group method administration upon the scorability of the records. *Rorschach Res. Exch.*, 1943, 7, 54–69.

RORSCHACH "ORGANIC SIGNS" AND INTELLIGENCE LEVEL[1]

WALLACE C. DIERS
AND
CLINTON C. BROWN
VETERANS ADMINISTRATION HOSPITAL, FORT THOMAS, KENTUCKY

IN the course of recent research into the intellectual and personality characteristics of a group of patients diagnosed with the neurological condition of multiple sclerosis, some interesting findings were obtained relevant to Hughes's "Signs of Intracranial Pathology" [3]. These findings led to the problem to be investigated in this paper, that of the occurrence of certain types of Rorschach test behavior, supposedly peculiar to cases of organic brain damage, in patients without demonstrable organic involvement but of a low intellectual level. The method of investigation will involve a comparison of the occurrence of Hughes's "signs" in a group of multiple sclerosis (MS) patients of above-average intelligence with a group of MS patients of below-average intelligence, and a comparison of the latter with a control group of nonneurologic patients of a like intellectual level.

Multiple sclerosis is a neurological condition characterized by diffuse and spotty areas of demyelinated nerve tissue resulting from a chronic, residual and inflammatory process. The symptomatology is varied, most commonly including intention tremor, disturbances of gait, vision, coordination and speech as well as psychic involvements. According to Bing and Haymaker, " . . . [in MS] from autopsy to autopsy the topographic distribution of the lesions varies to such an extent that no two pathologic pictures are the same . . . the sclerotic plaques are scattered indiscriminately through the cord and brain" [1, p. 241]. The total picture is such as to suggest that the test performance of the MS patient might disclose the patterning suggested by Piotrosky, Rapaport, and Wechsler for the patient with organic brain damage.

In a previous study [2] the authors concluded that the "Deterioration Quotient" and the "Organic Brain Damage (OBD) signs" of the Wechsler-Bellevue Scale were either insufficiently sensitive to reveal OBD signs in the MS patient or that no cortical pathology existed. Particular attention was therefore given to an analysis of Rorschach test protocols of these patients in the hope that this technique, with its greater depth and sensitivity, would provide more conclusive evidence.

PROCEDURE

The total multiple sclerotic hospital population (N=31) was given complete diagnostic batteries. Six patients were eliminated from the study because of incomplete testing, leaving a total N of 25. The group was divided with respect to the group mean intelligence level (W-B IQ 102), into an above-mean group (Group H) and a below-mean group (Group L) of N=15 and N=10, respectively.

A control group was then selected from the nonneurologic patients of the hospital population. Care was taken in the selection to obtain a group whose mean intelligence closely approximated that of Group L, and in whom any suspicion of neurologic disorder leading to

[1]This study was conducted under the supervision of Thelma E. Brown, Ph.D., Chief Clinical Psychologist. Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the authors are the result of their own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

organic brain dysfunction had been ruled out by means of stringent neurologic, psychiatric and psychologic measures. This group, 11 in number, was designated Group C.

The authors assembled the Rorschach protocols and gathered the "Hughes signs" of intracranial pathology. The examiners' findings were reviewed by the authors in each case. Differences in scoring and interpretation were resolved by discussion. The Hughes signs were coded as indicated in Table 1.

TABLE 1
RORSCHACH DIAGNOSTIC SIGNS OF INTRACRANIAL PATHOLOGY (HUGHES)

Coded Letter	Sign	Wt. Score*
A	less than 15 *R*	1
B	less than 25 *R*	1
C	Not more than one *M*	2
D	C_n	1
E	Not more than one *FC*	1
F	Perseveration	2
G	*FM* > M	—1
H	Color shock	—2
I	Shading shock	—2
J	Confused succession	—1
K	Contamination	—1
L	Impotence	3
M	Perplexity	3
N	Automatic phrases	3

*Eighty-two per cent of brain pathology cases make a score of seven or more, according to Hughes [3].

TABLE 2
WECHSLER-BELLEVUE IQ's AND HUGHES'S SIGNS IN GROUPS H AND L

Group H				Group L			
Pt.	IQ	Hughes's Sign	Signs Wt. Score	Pt.	IQ	Hughes's Sign	Signs Wt. Score
Ha	132	BEI	4	Ha	98	AFLMN	12
St	124*	H	—2	Re	98	BCGJF	3
Ja	120	—	—	Ko	97	ACDEGLM	11
Re	120	—	—	Mi	95	BC	3
Br	117	—	—	Fo	94	GI	—
Pe	116	B	1	Do	93*	ACEF	6
Mi	114	I	—2	Ja	88	CM	5
Ko	107*	BEG	2	Mo	81	BE	2
Ba	105	G	—1	Ta	77	BF	3
Ru	105	J	—1	Id	76	ACDEF	7
Lu	105	AE	2				
Ru	104	ACEF	6				
La	104	G	—1				
Is	103	G	—1				
Au	102	I	—2				
N = 15		Total = 5		N = 10		Total = 52	
$\overline{M}$ = 111.6		$\overline{M}$ = .33		$\overline{M}$ = 88.9		$\overline{M}$ = 5.2	
		Sigma = 2.24				Sigma = 3.68	

*In these cases only the verbal subscale could be administered; verbal IQ's are therefore used for purposes of dichotomization.

TABLE 3
WECHSLER-BELLEVUE IQ's, HUGHES'S SIGNS, AND DIAGNOSTIC CLASSIFICATION, GROUP C

Pt.	IQ	Hughes's Sign	Signs Wt. Score	Diagnostic Classification
Ge	90	BHI	1	Hysteric
Wa	90	ACG	3	Arthritic
Ha	88	CEFI	5	Neurosis with potential psychosis
Jo	87	ACEL	7	Arthritic
Es	87	B	1	Bilateral amputee
De	87	BCF	5	Psychopathic with potential psychosis
Co	85	ACHM	6	Schizophrenia, paranoid
Da	84	BM	4	Hysteric
Bl	84	BCF	5	Bilateral amputee
Sa	79	AC	3	Somatic complaints
Ch	79	ACGM	6	Schizophrenia
N = 11		Total = 46		
$\overline{M}$ = 85.4		$\overline{M}$ = 4.18		
		Sigma = 1.90		

RESULTS

Tables 2 and 3 present the W-B IQ's and types of Hughes's signs found in each of the 25 MS patients comprising Groups H and L and in the 11 nonneurologic control patients comprising Group C.

Inspection of Tables 2 and 3 shows that not only is there a substantial difference between mean weighted signs on Groups H and L, but also that in Group H, 10 of the individuals showed a weighted score which was either zero or negative. In Group L, one individual had a score of zero and there were no negative scores. In Group C, all scores were plus. On the basis of the information provided in Hughes's article [3], it is not

TABLE 4
SUMMARY OF STATISTICAL FINDING FOR HUGHES'S SIGNS, GROUPS H, L, AND C

Group	SE_M	SE_{DM}	Diff M_1 & M_2	t
H	.620	1.44	4.87	3.382**
L	1.300	1.445	1.02	.706
C	.633			

**Significant at the 1 per cent level.

possible to evaluate the significance of a score less than the critical score of seven.

Statistical treatment yielded a highly significant difference between mean weighted scores on Groups H and L (above .01 probability level), and an insignificant difference between corresponding scores on Groups L and C. A summary of these findings is presented in Table 4.

The results summarized in Table 4 suggest an inverse relationship between IQ measured by the Wechsler Scale and the weighted Hughes score, independent of the factor of intracranial pathology. The difference between Groups H and L is apparently the factor of intelligence, not extent and duration of symptomatology, severity of disease, etc. Organic signs in multiple sclerosis should appear indiscriminately in the population, and should not distribute themselves according to the factor of intelligence, if the measures applied are valid ones.

Given a person of lower than average intelligence, and using the Hughes signs, it does not seem possible to determine whether he has declined from a previously higher level of function or whether his intellectual endowment was originally low.

It is suggested that further study of this problem should be made with a larger population similar to our control group.

SUMMARY

A group of 36 hospital patients were given complete diagnostic psychological testing and their records analyzed for signs of organic brain damage by the use of Hughes's signs of intracranial pathology. Twenty-five of these patients bore a diagnosis of multiple sclerosis; the control group of 11 carried miscellaneous and nonneurologic diagnoses. In the MS patients, dichotomized on the basis of group mean intelligence, the lower group exhibited a greater number of Hughes's signs and a higher weighted mean score to a statistically significant degree. No significant difference in mean Hughes's signs was found between the lower MS group and a control group. It is suggested that the Hughes signs inadequately discriminate between patients with intracranial pathology and patients of low original intellectual endowment. The validity of the sign approach is questioned and it is suggested that the use of the Hughes signs should be restricted to exploratory and experimental uses until their validity is established.

Received November 17, 1950.

REFERENCES

1. BING, R., AND HAYMAKER, W. *Textbook of nervous diseases.* St. Louis, Mo.: Mosby, 1949.
2. DIERS, W. C., AND BROWN, C. C. Psychometric patterns associated with multiple sclerosis. *Arch. Neurol. Psychiat., Chicago,* 1950, **63**, 760–767.
3. HUGHES, R. M. A factor analysis of Rorschach diagnostic signs. *Rorschach Res. Exch.,* 1948, **12**, 165–167.

PERFORMANCE ON THE HARROWER-ERICKSON MULTIPLE CHOICE TEST OF PATIENTS WITH SPINAL CORD INJURIES

HAROLD H. MOSAK

VETERANS ADMINISTRATION MENTAL HYGIENE UNIT, DENVER, COLORADO[1,2]

ONE of the most amazing medical feats after the recent war was achieved by medical officers of the armed forces and the Veterans Administration in keeping alive servicemen who had incurred spinal cord injuries which resulted in a paralysis of the lower portion of the body. Approximately three hundred of these paraplegic patients were being treated at Hines VA Hospital when this study was initiated in 1947. The very nature of the injury made psychological problems inevitable. Voluntary control of the portion of the body below the level of the transection of the spinal cord was impossible, the sexual functions were lost, and incontinence of the bowel and bladder created additional problems. Frustration, anxiety, and depression ruled supreme. Beyond the overt behaviorial manifestations, little knowledge was possessed of the patient's attitudes toward his condition, toward himself, and toward other people. It was with these conditions in mind that the writer performed the following experiment to investigate in what ways their outlook upon life differed from that of nonparaplegic veterans.

EXPERIMENTAL PROCEDURES

The Harrower-Erickson Multiple Choice Test of three hundred choices was administered to 25 paraplegic and 25 nonparaplegic veterans. The latter group, composed of volunteers from the American Veterans Committee, hereafter referred to as the AVC group,[3] was of nearly the same average age and socioeconomic status as the patient group. The hospital group, also a volunteer group, was small because the physical condition of many precluded their being tested. No quadriplegic patients (patients with injuries to the cervical region of the spinal cord preventing the use of both arms and legs) were included in the experimental group. The ten Rorschach blots were projected upon a screen and exposed for three minutes, during which time the subjects indicated three responses for each blot. Secondary responses were not permitted, so that the total number of responses was kept constant at 30 for each subject. Each response was scored numerically in the manner suggested by Harrower-Erickson [3], responses scored 1 through 5 constituting so-called "good" responses and including movement, good form, animal, good form-color, small detail, and space percepts. Responses scored 6 through 10 included poor form, vague percepts, pure color, card description, and card rejections. In addition to the numerical scoring, each response was scored in terms of determinants and content according to the system advanced by Beck [1]. With the exception of *Dd* responses it was impossible to judge the "location" of these percepts, for the subject did not indicate whether he saw the suggested percept as a whole or as a detail. This was especially true of such responses as "cloud," "smoke," and "animal skin." It remained

[1]Sponsored by the VA and published with the approval of the Chief Medical Director. The statements and conclusions published by the author are a result of his own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

[2]This study was completed while the writer was a clinical psychologist trainee at the VA Hospital, Hines, Illinois.

[3]The writer wishes to express his gratitude to those paraplegic patients who served as subjects, as well as to those members of Thomas Jefferson Chapter #243, American Veterans Committee, who so kindly volunteered to serve as the control group.

possible, therefore, to score for *M, CF, FC, Y, YF, FY, FV, F+, F—, Dd,* Sex, Blood, Dirt, *A, H, An, P, S,* Card Description, and Card Rejection.

TABLE 1

MEAN NUMBER OF RESPONSES FOR RORSCHACH LOCATION AND DETERMINANT CATEGORIES

Response	$M_{paraplegic}$	M_{AVC}	t
Dd	1.12	1.52	1.21
M	2.24	3.84	3.02
C	0.68	0.12	2.55
CF	2.88	1.44	3.43
FC	1.44	1.44	—
ΣC	4.62	2.34	3.74
Y	1.16	0.40	2.53
YF	0.84	0.56	1.08
FY	2.92	2.40	1.33
FV	0.12	0.16	0.29
F+	11.92	15.12	2.81
F—	3.16	1.64	2.41

RESULTS

When scored on a numerical basis, it immediately becomes apparent that the paraplegic group designated a larger number of "bad" responses. The mean for this group was 11.6 "bad" responses as against 6.6 for the AVC group, a difference significant at the 1 per cent level. Further scrutiny of the responses in terms of location, determinants, and content unearthed some of the factors responsible for this difference. Table 1 gives the *t*'s for location and determinant scorings of the test. We note that the differences for *M, CF,* and *F+* are significant at the 1 per cent level; *C, Y,* and *F—* at the 2 per cent level; and *ΣC* at the .1 per cent level of significance. The remaining differences were not statistically significant.

TABLE 2

MEAN NUMBER OF RESPONSES FOR MISCELLANEOUS RORSCHACH SCORING CATEGORIES

Response	$M_{paraplegic}$	M_{AVC}	t
A	10.24	11.28	0.78
H	3.64	6.48	3.19
An	4.28	2.28	2.04
Bl	0.56	0.12	2.20
Dirt	0.76	0.08	3.09
Fi	0.96	0.32	2.29
Sex	0.12	0.32	1.43
P	10.80	13.00	1.96
S	0.16	0.52	1.50
Card Des.	0.96	0.12	2.15
Card Rej.	1.52	2.12	0.94

Table 2 presents the *t*'s for content and miscellaneous scorings. Of these, six were statistically significant—*Bl, Fi, An,* and Card Description at the 5 per cent level, and Dirt and *H* at the 1 per cent level of significance.

DISCUSSION

What do these differences tell us qualitatively about the personality of the paraplegic veterans? Let us first examine the means for the movement responses and attempt to formulate a possible explanation for this difference. According to Table 1 the mean number of *M* responses for the patients was 2.24, while that for the AVC group was 3.84. The difference between means was at the 1 per cent level of significance. The movement responses, according to Beck [1] "are representatives of very deep wishes, innermost psychic activity." These responses, he continues, are "kinaesthetically determined." If we accept this hypothesis, then it is altogether reasonable to assume that a reduction in kinaesthetic stimulation, as occurs in paraplegia, would result in the production of fewer movement responses. To test the hypothesis that paraplegics tend to suppress movement in their projections, each subject in both groups was asked to draw a picture of a man and of a woman. While elaborate inferences may be made from the size, placement, facial expression, line pressure, continuity, and other elements of these drawings [2, 4, 5], the writer restricted himself to an examination of the expression of movement of the lower extremities in the figures drawn. Of the 21 paraplegics whose injuries did not prevent them from drawing, only 6 executed drawings in which the legs of *both* figures were functional, i.e., they stood with legs apart or walked. Of the AVC sample 12 of the 25 drew two figures in which "movement" was evident. The remainder of the drawings exhibited other differentiating characteristics—in the paraplegic group, for example:

1. The drawing of only the head for both man and woman,
2. The representation of only a head for the

man but a full-length drawing of the woman,

3. Profile or frontal views of both figures, the man with nonfunctional legs, the woman in a "movement" pose,

4. Both figures with rigid or flaccid, but nonfunctional, legs,
or

5. A complete rejection of the task when the subject was asked to complete the figure.

Even one figure which was apparently drawn by a mental defective demonstrates this realization of the inutility of the lower portion of the patient's body.

In the affective sphere many differences in emotional relationships with the environment are immediately evident. The writer employed the weights for *C, CF,* and *FC* customarily utilized in obtaining a $\sum C$ score, but it must be realized that since these responses were not spontaneously elicited, the $\sum C$ cannot be interpreted in the traditional Rorschach manner. Nevertheless, it can be utilized here as a relative index of the degree of affective stimulation and response. This difference, significant at the .1 per cent level, depicts the paraplegic as one who, in comparison with the nonparaplegic world, responds primarily to the emotional elements of his environment. Thus, while the nonparaplegic's strivings are most readily stimulated from within, the paraplegic is motivated most easily by external stimuli.

If we analyze the nature of the color responses, we can observe other striking differences. While many paraplegics are capable of mature, adult types of emotional rapport, the more prevalent emotional pattern is one of adolescent egocentricity. This is reflected in Table 1 in the difference in *CF* means. While this seems to be the most characteristic mode of emotional adjustment, elements of infantile, uncontrolled affective reactions are much in evidence, as indicated in Table 1 by the difference between means for the pure *C* response. We find manifestations of this potential explosiveness both in the number of pure color responses and in the number of Fire responses. The latter include all such responses as "smoke," "an erupting volcano," and "an exploding firecracker."

In his treatment of the *Y* responses, Beck [1] refers to them as expressions of an "anergic" state and a "dysphoric" mood in the individual, amounting in the pure *Y* responses to passivity or total absence of activity. A priori we should expect to discover both conditions in the paraplegic population, and this is indeed the case. More *Y* responses were elicited from this group, and of these the largest difference is manifest in the pure *Y* responses, the indicators of what Beck terms a "washed-out"feeling.

An inspection of the content of the responses yields further illuminating aspects of the paraplegic's adjustment. We have previously discussed the importance of the Fire content in relationship to the patient's affective experiences. Human responses reflect a need for relationships (not necessarily healthy ones, however) with other people. The AVC group exhibits a greater gregariousness, a greater facility for forming associations with other people. Table 2 gives the means for the AVC and paraplegic groups at 6.48 and 3.64, respectively, the difference being significant at the 1 per cent level. There is in this statement no implication, however, that paraplegics are less able to conform to group standards or to adhere to the modes of thinking of the social milieu. Comparison of the *A* and *P* categories for both groups shows no significant difference and confirms this essential respect for social convention and cliché.

Specific personal needs crop out in the content designations. The Multiple Choice Test records of the paraplegic are rife with Anatomy, Blood, and Dirt responses, the latter category embracing such responses as "something squashed," "a black mess," and "dirty water." These responses likely reflect their preoccupation with their injuries, with the incessant treatment, and with the "messiness" of their conditions, especially with respect to incontinence.

With respect to intelligence and intellectual efficiency, little, unfortunately, can be deduced because of the very nature of the testing instrument used. Since many of the responses were not scorable in terms of *W* and *D,* we could not form any opinions regarding Approach and Sequence, but we do know that the groups did not differ in attention to small

details. We might expect that intellectual efficiency would remain unimpaired, and the writer believes that this conjecture is not contradicted by the data. Why then should there exist a statistically significant difference in the $F+$ response? We must recall that this is the mean of the $F+$, not the mean of the $F+\%$, and consequently not a measure of intellectual integrity but rather an indication of interest in the formal aspects of the blots. By this we mean that the construction of the test can be made to account for this discrepancy; since the number of responses was limited for each individual, it is obvious that if the paraplegic group were primarily responsive to color, they could not simultaneously select $F+$ percepts. This appears to be the most reasonable explanation for our data. Yet this does not explain the difference in the $F-$ category. Examination of the $F-$ responses, however, reveals that most (69 per cent) of these temporary breakdowns occur in conjunction with Anatomy responses. Thus, the pressing personal needs related to concern over health are apparently responsible for the distortion of reality and the corresponding difference in the $F-$ means.

The evasions constitute the last group of responses we shall consider. These include the Card Descriptions and the Card Rejections, and if we group these categories together, little difference in evasive behavior can be said to exist. It is interesting to note in this respect, however, that none of the paraplegics disguised their names, although the instructions permitted them to do so, while seven of the AVC group did. Nevertheless, the paraplegic sample, by virtue of its extreme stimulation by color, apparently selected a "Just colors" response in preference to a "Nothing at All" response whenever it had an opportunity to do so.

SUMMARY

The Harrower-Erickson Multiple Choice Test was administered to volunteer groups of 25 paraplegic patients and 25 nonparaplegic veterans. Each of the responses was scored both numerically and in terms of determinants and content. Significant differences were discovered for the numerical scorings as well as for 13 of the scoring categories. The paraplegic may be characterized as one who is motivated with greater ease by stimuli external to himself and whose emotional makeup is primarily egocentric. However, the potential for mature emotional response as well as for explosive affect is still present. He is intensely preoccupied with his own condition and the "messiness" of his life, is constantly aware of the inutility of his legs, but knowing that protest will not alter his condition, has become passive in his relationships. The facility for association with other people has decreased, but he still observes group mores.

Received December 18, 1950.

REFERENCES

1. BECK, S. J. *The Rorschach test.* New York: Grune & Stratton, 1946.
2. BUCK, J. N. The H-T-P technique: A qualitative and quantitative scoring manual. *J. clin. Psychol.*, 1948, **4**, 317–396.
3. HARROWER-ERICKSON, MOLLY R., AND STEINER, M. E. *Large scale Rorschach techniques.* Springfield, Ill.: Charles C Thomas, 1945.
4. MACHOVER, K. *Personality projection in the drawing of the human figure.* Springfield, Ill.: Charles C Thomas, 1949.
5. WOLFF, W. *The personality of the pre-school child.* New York: Grune & Stratton, 1946.

MMPI PROFILES IN PATIENTS WITH LOW-BACK PAIN[1]

LEO J. HANVIK
WASHBURN MEMORIAL CLINIC, MINNEAPOLIS, MINNESOTA

ALTHOUGH the literature relating to the subject of low-back pain is extensive, perusal of this literature has indicated that the standard techniques of clinical psychology have not been widely used to aid in evaluating the functional elements in patients with this syndrome. The following study is an attempt to see if patients whose complaint of low-back pain is deemed psychogenic in origin present a characteristically different picture psychologically from patients with like physical symptoms but whose symptoms are known to be produced by organic disease. The patients studied here were all patients at the Minneapolis Veterans Administration Hospital whose primary reason for hospitalization was low-back pain.

PROBLEM

Problem Statement. This investigation was undertaken to determine if, and to what extent, the Minnesota Multiphasic Personality Inventory can be used in differentiating between patients with "functional" low-back pain and patients who likewise have low-back pain but in whom there is evidence of the presence of organic disease.

Back Pain Defined. In the selection of cases, two criteria were followed:

1. Back pain must have been the primary reason for hospitalization.

2. Pain originated or appeared to originate in the back, in the spinal column or in close proximity thereto, from the lowest thoracic vertebrae to the coccyx. There may have been, and frequently was, other pain, (radiated or associated), but the source and basis of pain was deemed to be the lower back.

Description of Patient Sample. The subjects were all male inpatients admitted to the above-mentioned hospital with the complaint of back pain as defined above. The group actually constituted two clinical groups:

1. Thirty proved "organic" cases of low-back pain; all of these were patients with protruded intervertebral disc, and in all but two of these the evidence of the presence of disc was surgical, i.e., a disc was removed at operation. In the other two cases the diagnosis of disc was based on X-ray evidence (spinogram) and characteristic history, plus evidence of correctly distributed pain.

2. Thirty cases with no clear-cut organic findings, i.e., the general physical and neurological examinations were essentially negative.

The groups were equated for age (within five years), socioeconomic group (1940 Census classification), marital status, intelligence (within Stanford-Binet categories, since the measure of intelligence used was the Stanford-Binet Vocabulary), and race (all white).

Method of Analysis of Test Results. The following methods were used with the MMPI to test the hypothesis that the two clinical groups were not significantly different with respect to their test scores and patterns:

1. *t*-tests for the significance of differences between mean *T*-scores of the two groups on the MMPI scales.

[1]The material contained in this paper is abstracted from a Ph.D. thesis submitted to the Graduate Faculty of the University of Minnesota. The paper is published with the permission of the Chief Medical Director, Department of Medicine and Surgery, Veterans Administration, who assumes no responsibility for the opinions expressed or the conclusions drawn by the author. Special thanks are due to Dr. Joe R. Brown, M.D., formerly Chief, Neuropsychiatric Service, Minneapolis Veterans Administration Hospital, now of the Mayo Clinic, Rochester, Minnesota, for his assistance in this investigation.

2. Pattern analysis of the MMPI coded profiles for the two groups.

3. Test of the ability of experienced clinicians to separate the MMPI profiles into the correct groups in a manner significantly different from chance.

RESULTS

How the MMPI Differentiates the Two Groups. An inspection of Figure 1 and Table 1 will show that certain marked differences exist between the MMPI profiles of the two groups of patients.

Figure 1 illustrates that the "composite" profile for the 30 functional patients shows a fairly clear-cut "conversion-V" or "depressive valley" [1] configuration; i.e., Hypochondriasis and Hysteria are both elevated and Depression is low, relatively. This V-shaped pattern has aptly been termed the "conversion-V" because it is so frequently observed to occur in the records of conversion hysteria patients and because it so clearly portrays the essential elements of the hysterical adjustment where there are conversion features: e.g., the existence of physical symptoms along with indifference (low *D*) or apparent dissociation of the affective reactions. This type of record quotes the patient as saying, in effect, "I have numerous bodily complaints but I am relatively unworried, not depressed." The composite profile for the organics, it can be seen, does not show this pattern but, instead, is somewhat of a straight line, with *Hs, D,* and *Hy* approximately equal.

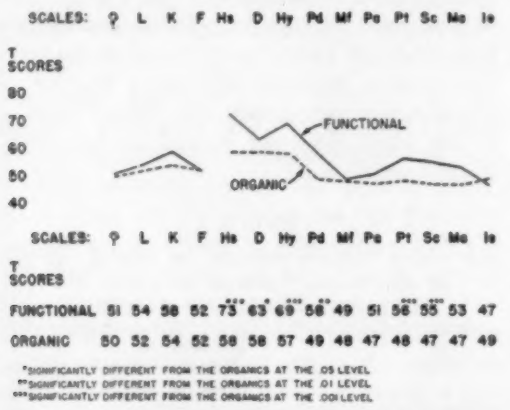
SCALES: ? L K F Hs D Hy Pd Mf Pa Pt Sc Ma Is
T SCORES
80
70
60
50
40
FUNCTIONAL
ORGANIC
SCALES: ? L K F Hs D Hy Pd Mf Pa Pt Sc Ma Is
T SCORES
FUNCTIONAL 51 54 58 52 73* 63* 69** 56* 49 51 56** 55** 53 47
ORGANIC 50 52 54 52 58 58 57 49 48 47 48 47 47 49
*SIGNIFICANTLY DIFFERENT FROM THE ORGANICS AT THE .05 LEVEL
**SIGNIFICANTLY DIFFERENT FROM THE ORGANICS AT THE .01 LEVEL
***SIGNIFICANTLY DIFFERENT FROM THE ORGANICS AT THE .001 LEVEL

FIG. 1. Composite MMPI profiles plotted from mean *T*-scores for the 30 functional low-back pain cases and for the 30 organic low-back pain cases.

TABLE 1

MMPI MEAN T-SCORES, VARIANCES, AND *t*'s FOR THE 30 HOSPITALIZED "ORGANIC" BACKACHE CASES AND THE 30 HOSPITALIZED "FUNCTIONAL" BACKACHE CASES

Scale	Organic Means	Organic Variances	Functional Means	Functional Variances	t
K	53.87	54.45	58.53	95.93	2.04
Hs	57.03	139.36	73.29	122.25	6.23***
D	57.50	66.40	62.83	97.73	2.2*
Hy	56.60	88.96	69.36	74.76	8.2***
Pd	49.00	128.94	57.50	62.07	3.32**
Mf	48.17	54.80	48.60	73.77	0.36
Pa	46.90	48.64	50.73	100.99	1.68
Pt	48.47	60.45	56.40	61.10	3.90***
Se	47.47	61.71	55.23	41.49	4.11***
Ma	47.49	55.57	52.96	85.66	1.30
Is	48.53	57.57	46.40	55.99	0.20

*Significant at the .05 level.
**Significant at the .01 level.
***Significant at the .001 level.

When mean *T*-scores of the two groups are compared (see Table 1), a number of significant differences are found: the mean *T*-scores on *Hs, Hy, Pt, Sc,* and *Pd* show very significant differences; the groups are also significantly different on the *D* scale, though the difference is somewhat less pronounced than on the five scales previously mentioned. The functional group, it is noted, has the higher scores on all these scales.

Analysis of the profile codes for the two groups, though it could not be carried out in statistical terms, also revealed differences: for example, *D* moves downward in importance as one progresses from the organic through the functional group of cases. *Hy* moves to the fore in rank as one progresses in the same direction.

Profile-Sorting Experiment.[2] After it was established that statistically significant differences exist between the MMPI scores for the two clinical groups, it was thought desirable to get an idea of the extent to which these differences could be used by clinicians in differentiating the profiles of the two groups by inspection. A profile-sorting experiment was therefore carried out with the aid of four clinical psychologists, all experienced in the clin-

[2]The author acknowledges indebtedness to Dr. Paul E. Meehl for his assistance in this experiment.

ical use of the MMPI. Only one judge had any knowledge at the outset that the sorting had anything to do with the low-back study, and that judge did only the "second sorting," to be explained subsequently. The other three judges sorted the profiles twice.

For the first sorting the judges were instructed in this manner:

These are curves of males (N = 60) from the ages of 20 to 55, median age 35. Sort them into piles of diagnostic categories as follows:

1. Essentially normal
2. "Normal" with physical disease
3. Simple adult maladjustment
4. Psychoneurosis:
   a) Reactive depression
   b) Hysteria
      i) Anxiety
      ii) Conversion
   c) Psychasthenia
   d) Neurasthenia
   e) Hypochondriasis
   f) Anxiety state
   g) Mixed
5. Manic-depressive:
   a) Manic
   b) Depressed
   c) Mixed
6. Schizophrenia:
   a) Simple
   b) Hebephrenic
   c) Catatonic
   d) Paranoid
7. Paranoid
8. Involutional Psychosis
   a) Melancholia
   b) Paranoid
9. Psychosis with psychopathic personality
10. Organic psychosis (e.g., paresis)
11. Psychopathic personality
    a) A-social, a-moral (Pd syndrome)
    b) Pathologic emotionality
    c) Pathologic sexuality

Record the case numbers after the diagnostic category in which you think it belongs.

After each of the three judges had performed this first sorting on the 60 MMPI profiles, he was given the following instructions for the second sorting:

These are all cases of veterans in the VA Hospital with the major complaint of low back pain. Of these 60 curves, there are 30 of each of two groups of cases: (1) male veterans with low back pain having no demonstrable organic finding upon thorough physical study; (2) male veterans with low back pain and a surgically confirmed prolapsed intervertebral disc. About half of these were tested as much as ten days after surgery, the rest were tested before surgery. Sort into two piles, one "functional" and one "disc". You may use or ignore your previous sorting as you see fit.

Examination of the results of the first sorting indicated that the psychotic categories (Nos. 5 to 10, inclusive, above) were largely ignored by the judges. That is, although the judges had no information whatever as to the character of the sample (other than the sex and age range), they were able to conclude correctly from the MMPI profiles that the patients represented thereby (*a*) were not psychotic, (*b*) had a sizable number of "normals" among them, and (*c*) had a large number of neurotics among them, particularly hysterical and hypochondriacal individuals. This fact is especially significant in view of the fact that the sorting was "blind," a condition which does not prevail in the actual clinical situation.

TABLE 2

RESULTS OF "SECOND SORTING" OF MMPI PROFILES OF 30 ORGANIC CASES OF BACKACHE AND 30 FUNCTIONAL CASES OF BACKACHE

	Per Cent Called Correctly	Chi Square	Probability
Judge A	77	9.187	<.01*
Judge B	73	6.909	<.01
Judge C	70	5.000	<.03
Judge D	57	0.536	Chance
Combined (Four Judges)	69	18.306	<.001

*Chi square for the .01 level is 6.635.

TABLE 3

CHI SQUARE USING ONLY CASES ON WHICH FOUR OR THREE OF THE JUDGES INDEPENDENTLY AGREE AS TO THE SORTING

			Contingency		
Agreement	*N**	%*	Chi Square	Co-efficient	Probability
Four Judges	29	83	12.461	.548	<.001
Three Judges	52	75	13.019	.447	<.001

**N* equals the number of cases which fit the criteria (i.e., the number of cases on which four or three judges independently agreed as to the category into which they should be sorted); "%" equals the per cent of correct hits (i.e., the per cent of *N* correctly called when we consider only those cases on which four or three judges independently agreed as to the sorting.)

Table 2 summarizes the results of the second sorting and indicates how well the four judges were able to sort the profiles into the categories "functional" and "organic." An inspection of the table reveals that all except one of the judges were able to sort the profiles significantly better than chance; and the total correct sortings for the four judges (including the one who failed to sort the profiles significantly better than chance) yielded a chi-square with a probability of .001.

Table 3 shows that when three or four judges independently agreed as to whether a given profile should be called "functional" or "organic," the accuracy of the sorting increased, so that the difference from chance was even more highly significant statistically.

We can conclude from the profile-sorting experiment that experienced clinicians can distinguish with good accuracy the MMPI profiles of patients with functional low-back pain from those of patients in the organic group. This is the case in spite of the fact that supplementary information and cues which are frequently available in the clinical setting were lacking in the present experiment.[3]

A word should perhaps be added here concerning individual differences in the ability to sort MMPI profiles. It is, of course, essential that an individual shall have had practice clinically in dealing with a considerable number and variety of MMPI profiles. The judges employed in the sorting experiment just described, for instance, have all had abundant experience in the use of the MMPI in diagnosis and have all viewed numerous profiles of neurotic, psychotic, and normal individuals. Even so, there appear to be differences in skill, and these differences have consistently exhibited themselves in a series of sorting tests like the one reported here.

For example, Judge D (see Table 2), whose sorting level was apparently "chance" in this experiment, was found to have been at or near the bottom in proficiency in at least two other sorting problems where his results were in competition with those of from three to five other judges. Moreover, this is true notwithstanding the fact that Judge D has had at least as much clinical experience with the MMPI as any judge with whom he has been "in competition." There are other instances known to the author in which certain experienced judges consistently ranked low or high in sorting in competition with others. It would thus seem that sorting, or better, "evaluation" of MMPI profiles is a clinical activity or skill requiring good judgment sharpened by experience and that, even among psychologists who have had approximately the same amount of clinical experience with the test, there appear to be marked individual differences in the level of skill attained.

[3] As will be noted from the instructions given for the second sorting, some of the patients were tested before surgery and some afterward. By means of an item analysis and significance tests, the MMPI profiles of the group tested prior to surgery were compared with those of the group tested afterward in order to see if there were any differences. All tests indicated that there were no significant differences.

SUMMARY

A study was done to determine if the MMPI could be used in differentiating between cases of presumed psychogenic backache and backache cases in which there was definite evidence (surgical in all except two cases) of herniated intervertebral disc.

There were statistically significant differences between the two groups on the following six scales: *Hs, D, Hy, Pd, Pt,* and *Sc.* The functional patients scored higher on all of these scales. When the mean *T*-scores for all patients in the functional group were plotted, the resultant profile was neurotic in type, showing the "conversion-V" configuration, which features elevations on *Hs* and *Hy,* with *D* relatively low and a mild rise on *Pt.*

Profile-sorting experiments revealed that judges could distinguish the profiles of members of the two groups in a manner significantly better than chance.

Received November 13, 1950.

REFERENCES

1. GOUGH, H. E. Diagnostic patterns on the Minnesota Multiphasic Personality Inventory. *J. clin. Psychol.,* 1946, **2**, 23–37.
2. HANVIK, L. J. Some psychological dimensions of low back pain. Unpublished Doctor's thesis, Univ. of Minn., 1949.

NEW BOOKS AND TESTS

Books

BAUMAN, MARY K., AND HAYES, SAMUEL P. *A manual for the psychological examination of the adult blind.* New York: Psychological Corp., 1951. Pp. 58. $1.50.

The testing of the blind raises many problems for psychological examiners, which may be met by the use of techniques and resources that have been accumulated through years of experience. This useful handbook, prepared by the National Psychological Research Council for the Blind, opens with chapters on the problems of blindness, on special considerations in testing the blind, and on the case history of a blind person. The main part of the manual is an annotated inventory of techniques for measuring abilities, aptitudes, achievement, interests, and personality. There is a good chapter on the reporting of psychological examinations to meet the needs of rehabilitation counselors.

BURNHAM, PAUL S., AND PALMER, STUART H. *Counseling in personnel work: 1945–1949; an annotated bibliography.* 1313 E. 60th St., Chicago: Public Administration Service, Publication No. 105, 1951. Pp. 39. $1.00.

An indexed and briefly annotated bibliography of 423 books and articles dealing with all phases of counseling in public agencies: general, employee, occupational, rehabilitation, student, and veteran.

CASS, MARION T. *Speech habilitation in cerebral palsy.* New York: Columbia Univ. Press, 1951. Pp. xi + 212. $3.00.

The psychological and educational rehabilitation of children with cerebral palsy is offering a challenging new field for the efforts of psychologists. After a brief but good summary of the history and characteristics of cerebral palsy, the major part of this book is concerned with the teaching of speech. Detailed exercises are given for practice in all of the major speech sounds. Although no experimental evidence is offered to show that the author's system is better than any other, she does supply a technique which has been developed through broad practical experience, and which is described so explicitly that it can be followed by anyone with a general knowledge of speech correction.

NOTE: Some reviews in this issue were prepared by the Associate Editors, who may be identified by their initials. Unsigned reviews are by the Editor. —*L. F. S.*

ELLIS, ALBERT. *The folklore of sex.* New York: Charles Boni, 1951. Pp. 313. $5.00.

This is a popularly written report of a study of the liberal, conservative, or salacious nature of all sex references in some of "the most popular and widely representative mass media" of early 1950. The media studied are fiction and nonfiction best sellers, magazines (including religious and scientific journals), movies (except Westerns), plays and musicals, city newspapers, popular songs, and national radio and television programs. Twenty-seven of the 34 chapters cite sample attitudes on the various practices and conditions associated with sex relations and control. The other chapters present the summarized data and implications. The study shows, in the main, that the attitudes expressed are "conflicting, ambivalent, and confused" and further that the extent to which liberal, conservative, or salacious attitudes are expressed varies with the medium and the practice. The author implies from his data that "virtually every living American . . . is completely muddled-, mixed-, and messed-up in his sex views, feelings and acts." Certainly Dr. Ellis deserves credit for documenting the disparity that exists between our expressed sex attitudes, showing probable social causes, suggesting that existing controls often defeat their purpose, and providing a starting-point for the progressive controversy which the book no doubt will incite. However, the above implications seem to go beyond the data presented and will weaken his presentation. In the first place, this is not a study of the folk-lore of sex, but a presentation of the attitudes of groups heavily weighted with the urban, sophisticated, literate, and secular subcultures of the United States. Secondly, Dr. Ellis draws implications about the integration of conflicting sex attitudes and behavior (no doubt the most important problem) in individuals from verbalizations, largely fictional, designed for their consumption.—*F. McK.*

FORD, CLELLAN, S., AND BEACH, FRANK A. *Patterns of sexual behavior.* New York: Harper, 1951. Pp. viii + 307. $4.50.

In this book an anthropologist (Ford) and a comparative psychologist (Beach) combine to present sexual practices in the broad perspective of the behavior of 190 contemporary human societies and numerous animal species. The result is a broad extension of the Kinsey report, developed in much less ponderous fashion. It is sound, well written, and, needless to say, makes fascinating reading. This reviewer feels that no clinician can afford to overlook it; most of them will not want to.—*W. A. H.*

GOHEEN, HOWARD W., AND KAVRUCK, SAMUEL. *Selected references on test construction, mental test theory, and statistics, 1929–1949.* Washington: U.S. Government Printing Office, 1950. Pp. xii + 209. $1.50.

A classified and indexed bibliography, prepared in the Test Development Section of the U. S. Civil Service Commission. Although the original purpose of the list was oriented to tests for personnel selection, the wide coverage makes the references useful to all who are concerned with human appraisal.

HATHAWAY, STARKE R., AND MEEHL, PAUL E. *An atlas for the clinical use of the MMPI.* Minneapolis: Univ. of Minnesota Press, 1951. Pp. xliv + 799. $9.75.

The *Atlas* is a most unusual volume, consisting of 968 short histories of a wide variety of psychiatric cases, arranged by their coded profiles on the Minnesota Multiphasic Personality Inventory. Each case summary is mainly factual, describing the patient as he entered the hospital, his pertinent history, and how he appeared to the therapist or hospital staff. There is a minimum of dynamic interpretation. The cases were taken without reference to MMPI profiles, except for a few selected in the last stages of the preparation of the volume to provide examples of rare patterns. Although the authors, quite wisely, make no claim that the case studies validate the MMPI, reading the cases does suggest that the profiles group them according to their main behavioral trends. The aim of the *Atlas* is to provide supplementary clinical training for users of the Inventory. It assumes that the reader already has some knowledge of the MMPI and of the research underlying its construction. The *Atlas* makes readily available a wider variety of case pictures than most workers would meet in their clinical experience. The volume has shortcomings that the authors recognize, including the brevity of the case citations and their limited source, mostly from one hospital. As a new experiment in clinical training, however, the *Atlas* will be of wide interest.

HORROCKS, JOHN E. *The psychology of adolescence.* Boston: Houghton Mifflin, 1951. Pp. xxvi + 614. $4.50.

Although more than two decades have elapsed since Mead's pioneer work defined adolescence as primarily a cultural period, surprisingly few books on adolescence have focused on social adjustment. Horrocks does so. The general viewpoint is culture-oriented, and the first five chapters review the social relationships of adolescents in the family, in peer groups, and with contemporaries of the opposite sex. Intellectual, physical, and physiological phases of development are not neglected, however. All topics are treated with an emphasis on research findings, and the bibliographies are extensive. For the basic training of high school teachers and other workers with youth, the book as a whole compares very favorably with others in its field.

KASIUS, CORA. (Ed.) *Principles and techniques in social casework: Selected articles 1940–1950.* New York: Family Service Association of America, 1950. Pp. 433. $4.50.

A book of readings, reproducing 32 outstanding articles of the past decade from the journal *Social Casework.* The compilation inevitably invites comparison with a hypothetical collection that might have been made from a journal for clinical psychologists, such as the *Journal of Consulting Psychology.* To their credit, the social caseworkers are more concerned with the philosophy of their profession, and with methods for supervision and training. Social workers and psychologists seem to give about equal emphasis to techniques of practice. In research, their disparity is striking; not one of the articles in the volume is a report of a quantitative experimental study.

LEMERT, EDWIN M. *Social pathology.* New York: McGraw-Hill, 1951. Pp. ix + 459. $4.50.

This text on the ills of societies elaborates a consistent theory based on two main postulates, that behavior deviations are a function of culture conflict, and that a deviant person is differentiated by means of a concept of his role formed by interactions between him and the dominant culture. The theory is illustrated by data from seven areas of deviation: blindness, speech defects, radicalism, prostitution, crime, alcoholism, and mental disorder. Aspects of the theory are by no means unfamiliar. Psychological formulations concerning the significance of the self concept, and of the role of social interaction in defining a deviation, suggest parts of Lemert's theory, as does also the emphasis of semantics on the functions of naming in the development of social roles. Still, clinical psychologists will find much that is fresh in a sociologist's discussion of mental disorder as a social pathology.

LIBERMAN, SALLY. *A child's guide to a parent's mind.* New York: Henry Schuman, 1951. Pp. 145. $3.00.

Don't be misled by its subtle persiflage; this book is profound. Aphorisms, wit, satirical illustrations, style and format create an atmosphere that *is* mental hygiene. But this aura enshrouds contents which are masterpieces of epitome. Browse its *New Yorker* style, absorb its PTA wisdom, digest its projectivisms, relish its Gertrudesteinishness—then scan, scrutinize, contemplate, and try to penerate its pithy insight. There is pathos, sobriety, yet an empathically sensitive tenderness. Don't pass up this book, or be hoodwinked by its disarmingly simple candor. And this goes also for the tragi-comic sketches. The title, a miracle of understatement, conceals a masterful treatise on life's major profundities, disarmingly set forth in the guise of childish simplicity. —*E. A. D.*

MASLOW, A. H., AND MITTELMANN, BELA. *Principles of abnormal psychology.* (Rev. Ed.) New York: Harper, 1951. Pp. xv + 665. $5.00.

The revision of Maslow and Mittelmann's text augments some of the good features, and seems to have lost none of the qualities, that made the original edition so satisfactory. The basic viewpoint is unchanged, but there is even more emphasis on cultural factors in mental disorders. Material has been added on schizophrenia, psychosomatic disorders, and war neuroses. Three new chapters bring together quite effectively the conflicts and mechanisms found in infancy, in childhood, and in adolescence. The least improvement, when there might have been more, is in the area of therapy. The chapter on techniques of psychotherapy has been revised by adding short and not too well integrated sections on newer developments; the chapter on psychoanalysis is substantially unchanged. Those who liked the first edition for its dynamic and social approach will like the revision also.

PEAR, T. H. (Ed.) *Psychological factors of peace and war.* New York: Philosophical Library, 1950. Pp. 262. $4.75.

This collection of essays, prepared for the United Nations Association, dispassionately considers war as a phenomenon of cultural behavior. Its objective attitude and its reliance on experimental observations distinguish it from many other well-intentioned but essentially flabby books in its field. Among the highlights are Eysenck's survey of social attitudes toward war and agressiveness, Himmelweit's review of experimental work on frustration and aggression, and Richardson's statistical analysis of wars. While the volume's primary concern is the social psychology of war and peace, many of the chapters contain major contributions to the study of attitudes and of personality.

RAMSEY, GLENN V. *Factors in the sex life of 291 boys.* Madison, N. J.: 82 Samson Ave., 1950. Pp. vii + 107. $2.00.

This is a lithoprint reproduction of the essentials of Ramsey's 1941 doctoral dissertation, republished because of continued demand. The study was supervised by Dr. A. C. Kinsey, and the design and results are similar to those published by him. The monograph includes data on the acquisition of sex information and the nature of sexual maturational problems. The age of 95 per cent of the subjects was from 12 to 18. There were two 100 per cent groups from the 7th and 8th grades, and other subjects were volunteers from boys' clubs and represented about 60 to 70 per cent of their group. The data from these samples were pooled. The frequencies of all types of sexual outlets are of the same relatively high magnitude as those in Kinsey's volume. Ramsey deserves praise for his courage and ingenuity. In view of his penetrating interrogation, it is surprising that he did not meet greater adult resistance to the study, despite the fact that it was well planned as a sex-education endeavor.—*F. McK.*

SHOCK, NATHAN W. *Trends in gerontology.* Stanford, Calif.: Stanford Univ. Press, 1951. Pp. ix + 153. $2.50.

SHOCK, NATHAN W. *A classified bibliography of gerontology and geriatrics.* Stanford, Calif.: Stanford Univ. Press, 1951. Pp. xxvii + 599. $15.00.

These companion volumes, one small and one very large, are the first fruits of a foundation-subsidized survey of the problems of aging. *Trends* ably summarizes the present state of gerontology. with data on population changes, employment, health, community programs, and education. Emphasis is placed on research in progress, on methods for furthering research, and on the need for a coordinating Institute of Gerontology. This brief book is as readable as it is informative, and will serve the needs of psychologists who want an overview of the field. The *Bibliography,* a major by-product of the survey, lists 18,036 titles from the world literature on all aspects of aging: biological, medical, psychological, social, and economic. The classifications are detailed and the volume is well indexed, but the entries are not annotated. References are given to published abstracts as well as to the original articles and books. Supplements to the bibliography are appearing in the *Journal of Gerontology,* to keep the material current. Research workers will find it indispensable.

THORNDIKE, ROBERT L. (Chm.) *Proceedings of the invitational conference on testing problems, 1950.* Princeton, N. J.: Educational Testing Service, 1951. Pp. 117. $1.00.

The 1950 conference sponsored by the ETS con-

sisted of two panel discussions and an invited address. Clinical psychologists will be especially interested in the papers on the validation of professional aptitude batteries, including those for law, dentistry, and medicine. L. L. Thurstone's address on creative talent is a striking contribution to that elusive topic.

VERNON, PHILIP. *The structure of human abilities.* New York: Wiley, 1950. Pp. 160. $2.75.

This rather short book contains a full survey of the results and implications of factorial analysis, with a brief nontechnical explanation of the methods. The author states that, "I assume only that the reader has had an elementary course in psychology and knows what an intelligence test and a correlation coefficient are," and for the most part this is all that is necessary to follow the text. The data are well organized and clearly presented. This is more than a survey which covers the American as well as the British work comprehensively; it is also a critical review with an appendix presenting persuasively the argument for the superiority of general plus group-factor solutions over multiple-factor solutions. This should be an extremely helpful book.—*A. R.*

WITTY, PAUL. (Ed.) *The gifted child.* Boston: Heath, 1951. Pp. xii + 338. $4.00.

In a field in which there is so much controversy, it is refreshing to find a book that faces complications squarely without oversimplification or bias. Sponsored by the American Association for Gifted Children, the volume was written by recognized leaders in its field. One is impressed with the quality of their contributions. All authors agree on the urgent need for utilizing the potential strengths of gifted children, but they do not agree on the means. Opposing viewpoints are expressed, and solutions are not offered as final ones. Rather, the reader is stimulated by the presentation of the issues and of some of the differing ways in which they are now met. An extensive, well-annotated bibliography and a good index increase the availability of the material. The book is well edited, and will appeal to psychologists, educators, and lay readers alike.—*B. M. L.*

Training of psychological counselors. Report of a conference held at Ann Arbor, 1949 and 1950. Ann Arbor, Mich.: Univ. of Michigan Press, 1951. Pp. 29.

Aided by a grant from the National Institute of Mental Health, a panel of twelve psychologists met in 1949 and 1950 to consider the work and training of psychological counselors. Their report is so important, and already so condensed in its present form, that it is hardly possible to abbreviate it further in a review. The conference considered and clarified, although it did not presume to solve completely, the definition of psychological counseling, and the differences between a counselor and a clinical psychologist. To meet realistically the needs for services and the resources available to support them, the conference advocated the training of psychological counselors at less than the doctoral level, in one-year and two-year graduate programs. At both levels, training emphasis was given to personality theory, appraisal, and counseling techniques. Practicum training was regarded as essential: " . . . no one should attempt to counsel independently until he has had a period of closely supervised experience." All psychologists concerned with professional training should give careful attention to this well-prepared and thought-provoking document.

CORRECTION

A typographical error occurred in the article, "Norms for 'Shock' in the Rorschach," by Herbert Sanderson, which appeared in the April, 1951, number of this *Journal.* On page 128, in Table 1, the mean reaction time for Card V should read 9.66, instead of 19.66.

NEUROLOGY & PSYCHIATRY

Section VIII of Excerpta Medica

The Neurology Section includes abstracts of all articles on neurological surgery in addition to general clinical neuro-physiology.

The section on Psychiatry covers the topics usually contained within the meaning of the word. In addition, the following special branches are dealt with fully:

* **Psychoanalysis, individual and analytical psychology**
* **Sexology, criminology, alcoholism, and drug addiction as related to psychiatry**
* **Social and industrial psychology and psychiatry, vocational guidance and mental testing**
* **Psychometrics**
* **Personality testing and the Rorschach test**
* **Heredity and statistical studies as they apply to psychiatry**
* **Mental defect and epilepsy**

Annual volume of 1000 pages, including an index classified by author and subject.

Subscription price $22.50 per annual volume

THE WILLIAMS & WILKINS COMPANY

Agents for the United States, Canada and Central America

Mt. Royal and Guilford Aves. Baltimore 2, Maryland

Please enter my subscription to Section VIII, NEUROLOGY AND PSYCHIATRY, of EXCERPTA MEDICA. I enclose $22.50.

Name.....

Address.....

.....

JOURNAL OF CONSULTING PSYCHOLOGY
AVAILABLE BACK ISSUES

The issues listed below are available for sale

YEAR	VOLUME	JAN FEB	MAR APR	MAY JUN	JUL AUG	SEP OCT	NOV DEC	PRICE PER NUMBER	PRICE PER VOLUME
1937	1	1	2	3	4	5	6	$.60	$3.00
1938	2	1	2	3	4	5	6	$.60	$3.00
1939	3	1	2	3	4	5	6	$.60	$3.00
1940	4	1	2	3	4	5	6	$.60	$3.00
1941	5	1	2	3	4	5	6	$.60	$3.00
1942	6	1	–	3	–	–	6	$.60	$1.80
1943	7	–	2	3	4	5	6	$.60	$3.00
1944	8	1	–	3	4	5	6	$.60	$3.00
1945	9	1	2	3	4	5	6	$.60	$3.00
1946	10	1	2	3	4	5	–	$.60	$3.00
1947	11	–	–	3	4	5	6	$.60	$2.40
1948	12	1	2	3	4	5	6	$1.00	$5.00
1949	13	1	2	3	4	5	6	$1.00	$5.00
1950	14	1	2	3	4	5	6	$1.00	$5.00
1951	15	By Subscription, $5.00							

Price for Volumes 1 through 14 $46.20

Table based on inventory of January 1, 1951

The Journal of Consulting Psychology is of particular interest to clinical psychologists, psychiatrists, school psychologists, and persons who are engaged in counseling and guidance work.

The price per volume of the Journal of Consulting Psychology was $3.00 between 1937 and 1947. It is now $5.00 per volume and $1.00 per issue. For foreign orders, $.25 per volume should be added. Address orders to:

AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 Massachusetts Ave. N. W.
Washington 5, D. C.